M 332 Unit 0

THE OPEN UNIVERSITY

Complex Analysis
Mathematics: A Third Level Course

Unit 0 Real Analysis Review

COMPLEX ANALYSIS

THE OPEN UNIVERSITY

Mathematics: A Third Level Course

Complex Analysis Unit 0

Real Analysis Review

Prepared by David Brannan

The Open University Press

The Open University Press
Walton Hall, Milton Keynes
MK7 6AA

First published 1981. Reprinted 1985. Second edition 1988.

Copyright © 1981, 1985, 1988 The Open University

All rights reserved. No part of this work may be reproduced in any form, by mimeograph or any other means, without permission in writing from the publisher.

Designed by the Graphic Design Group of the Open University.

Printed in Great Britain by
Speedlith Photo Litho Ltd, Longford Trading Estate, Manchester M32 0JT.

ISBN 0 335 05565 6

This text forms part of an Open University Third Level Course.

For general availability of supporting material referred to in this text, please write to Open University Educational Enterprises Limited, 12 Cofferidge Close, Stony Stratford, Milton Keynes, MK11 1BY, Great Britain.

Further information on Open University courses may be obtained from the Admissions Office, The Open University, PO Box 48, Walton Hall, Milton Keynes, MK7 6AB.

2.1

CONTENTS

Page

0 Introduction 5

1 R—The Real Number System 6
- 1.1 Inequalities 6
- 1.2 The Least Upper Bound Property of R . . . 11
- 1.3 The Archimedean Property of R 17

2 Sequences 19
- 2.1 Convergence of Sequences 19
- 2.2 Some Properties of Convergent Sequences . . . 26
- 2.3 Nested Intervals Theorem 28
- 2.4 Sequences of Functions 30

3 Limits of Functions, Continuity 33
- 3.1 Limits 33
- 3.2 More Limits 37
- 3.3 Continuity 41
- 3.4 Some Properties of Continuous Functions . . . 45
- 3.5 Uniform Continuity 48

4 Differentiation 49
- 4.1 The Derivative 49
- 4.2 Properties of the Derivative 52

5 Series and Power Series 55
- 5.1 Convergence and Absolute Convergence of Series . . 55
- 5.2 Power Series 57
- 5.3 Partial Fractions 59

6 Integration 63
- 6.1 Some Inequalities for Integrals 63
- 6.2 Another Use for Partial Fractions 64

7 Partial Derivatives 66

0 INTRODUCTION

Complex Analysis is the study of properties such as continuity, differentiability and integrability of functions whose domain and codomain are sets of complex numbers. A knowledge of Real Analysis illuminates Complex Analysis in two ways. It provides a repertoire of methods and results that may be applied directly or indirectly to Complex Analysis. It is also a standard of comparison: one of the fascinating things about Complex Analysis is the extent to which it differs from Real Analysis.

This is why we study Real Analysis in this unit.

There is more than one way of teaching any mathematical topic. As it happens, the approach to teaching analysis, based on convergent sequences in **R** in *M203, Introduction to Pure Mathematics*, is superficially rather different from, though logically equivalent to, the so-called "ε-δ approach" used in *M332*. (The equivalence of the two approaches was explained in *M203*, *Analysis Block B, Unit 5, Proofs*.) Thus, for students who have previously studied *M203*, this unit has the role of a conversion course (so far as the *methods* of analysis are concerned) as well as a review of *results*.

It is absolutely essential that you master the "ε-δ approach" before you start work on the rest of the course. In particular, when you check your mastery of the contents of this unit by tackling the Self-assessment Questions, *do not* tackle them using the "sequences approach" of *M203*, but instead use the "ε-δ methods" outlined in this unit.

If you learnt the rigorous approach to Calculus, that we nowadays call Analysis, in some other course, you will already have met the ε-δ approach. Nevertheless, we suggest that you read the material in this unit carefully to refresh your memory; who knows, you may meet something new! Also, try the Self-assessment Questions to make sure you really do remember your previous studies!

None of the material in *Unit 0* will be examined specifically in the Final Examination of the course; although, naturally, some knowledge of the contents of *Unit 0* will be essential to answer questions on topics covered in the remainder of the units.

Background Reading

If you feel unsure about any part of real analysis, for example about anything covered in this unit, probably the best reference book you could consult is *Calculus* by M. Spivak, published by W. A. Benjamin/Addison-Wesley (paperback edition, 1973). This is referred to as: **Spivak**.

Study Guide

You should read the sections of this unit in their natural order.

Before starting work on *Unit 1*, you should have read and fully understood the following sections of *Unit 0*: 1.1–1.3, 2.1–2.3, 3.1–3.4, 4.1 and 4.2, 5.1–5.3, and 6.2. These sections will be included in the continuous assessment component of the course.

Although the remaining sections of *Unit 0* (Sections 2.4, 3.5, 6.1 and 7) are intended to be read rather quickly at this stage, it is, of course, essential that you read each of them before the first unit in the course that depends on them. This means that you should have mastered Section 3.5 before starting work on *Unit 2*, Section 7 before *Unit 3*, Section 6.1 before *Unit 4*, and Section 2.4 before *Unit 11*.

There is a television programme associated with this unit.

1 R—THE REAL NUMBER SYSTEM

In this section we look at those properties of the real number system, **R**, that will be of particular importance in the rest of the course. In particular, therefore, we shall not give a logical construction of the real numbers, but simply assume that they are "there", and that our business for the moment is to remind you how they can be used.

In *M203* the set of real numbers
denoted by ℝ. In *M332* we use **R**

We shall often be using the pictorial representation of **R** as "the real number line"; in other words, the identification of each real number with some point on a straight line. Similarly, we shall often use the representation of the complex number system, **C**, as "the complex plane". The differences between real analysis and complex analysis are the differences between the properties of functions whose domains are **R** and **C**. These are quite striking, as you will see throughout the course.

M203 uses ℂ to denote the se
complex numbers.

1.1 Inequalities

We mention without further explanation the (usual) operations of addition and multiplication on elements of **R**. As inequalities will play a key role in *M332*, recall that **R** also has a rather special property called *order*, which can be expressed in terms of the symbol $\leqslant$; this is that for any pair x, y in **R**, either $x \leqslant y$ or $y \leqslant x$. Then, for example, if $x \leqslant y$, we have:

(i) $x + z \leqslant y + z$ for any z in **R**; and
(ii) $xz \leqslant yz$ for any z in **R** with $z \geqslant 0$.

$\leqslant$ means "less than or equal to"

Example

Find the set S of all $x \in$ **R** such that

$$x^2 - x - 2 \leqslant 0.$$

Solution

First we factorize the polynomial in x to rewrite the inequality defining S as

$$(x+1)(x-2) \leqslant 0.$$

There are various ways of analysing this inequality; we present one below.

The factor $x + 1$ is zero when $x = -1$, negative if $x < -1$, and positive when $x > -1$; similarly, the factor $x - 2$ is zero when $x = 2$, negative if $x < 2$ and positive if $x > 2$. To find the sign of the product $(x + 1)(x - 2)$, it is convenient to put this information in tabular form.

value of x	$x < -1$	-1	$-1 < x < 2$	2	$x > 2$
sign of $x + 1$	$-$	0	+	+	+
sign of $x - 2$	$-$	$-$	$-$	0	+
sign of product	+	0	$-$	0	+

This shows at once that

$$S = \{x \in \mathbf{R}\colon -1 \leqslant x \leqslant 2\}.$$

Self-assessment Question

Determine the set S of all $x \in \mathbf{R}$ such that

$$\frac{x+2}{x^2 - 3x} < 0.$$

Solution

Obviously $x = 0, 3$ are not valid values of x, since then the denominator vanishes.

We can rewrite the inequality defining S in the convenient form

$$\frac{x+2}{x(x-3)} < 0,$$

and then use a table of signs to find S.

value of x	$x < -2$	-2	$-2 < x < 0$	0	$0 < x < 3$	3	$x > 3$
sign of $x + 2$	$-$	0	+	+	+	+	+
sign of x	$-$	$-$	$-$	0	+	+	+
sign of $x - 3$	$-$	$-$	$-$	$-$	$-$	0	+
sign of $\dfrac{x+2}{x(x-3)}$	$-$	0	+		$-$		+

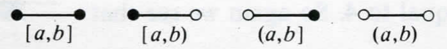
Hence

$$S = \{x \in \mathbf{R}\colon x < -2\} \cup \{x \in \mathbf{R}\colon 0 < x < 3\}.$$

Next we mention the *M332* notation for intervals in $\mathbf{R}$. The symbol $[a, b]$ is just short-hand notation for the set (closed interval)

$$\{x\colon a \leqslant x \leqslant b\},$$

a square bracket meaning that the corresponding end-point belongs to the set under consideration. If one or both square brackets is replaced by a round bracket, this means that that particular end-point of the interval does not belong to the specified set; thus, $[a, b) = \{x\colon a \leqslant x < b\}$, $(a, b] = \{x\colon a < x \leqslant b\}$ and $(a, b) = \{x\colon a < x < b\}$.



We allow the left end-point (right end-point) of an interval to be $-\infty$ $(+\infty)$ if appropriate. For example,

$$(-\infty, 2] = \{x \in \mathbf{R}\colon x \leqslant 2\} \quad \text{and} \quad (-\infty, \infty) = \mathbf{R}.$$

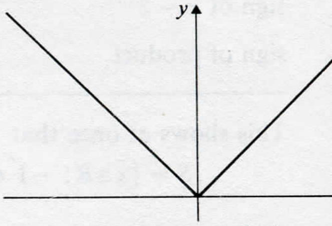
M203 used reversed square brackets, not round brackets.

The complex number system **C** has operations of addition and multiplication completely analogous to those of **R**. However, as you will see on page 13 of *Unit 1* it is *not* possible to define an order operation on elements of **C**. This means, for example, that any statement of the form

$$z_1 \leqslant z_2 \qquad (\text{where } z_1, z_2 \in \mathbf{C})$$

simply makes no sense (so be careful!).

Associated with the idea of order is the **modulus function**, which will play a key role in the rest of the unit and, indeed, in the whole course.



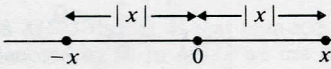
graph of $y = |x|$

For all $x \in \mathbf{R}$, the **modulus** of x, which we denote by the symbol $|x|$, is the non-negative real number such that

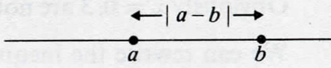
$$|x| = \begin{cases} x, & \text{if } x \geqslant 0, \\ -x, & \text{if } x < 0. \end{cases}$$

For example, $|0| = 0$, $|3\frac{1}{2}| = 3\frac{1}{2}$ and $|-5\frac{1}{2}| = 5\frac{1}{2}$.

Rather loosely, we can express the formula for $|x|$ and the fact that $|-x| = |x|$ for all $x \in \mathbf{R}$ as follows: $|x|$ is the magnitude of x without its sign. We can also regard $|x|$ as the distance of the number x from 0 on the real number line. Indeed, in the same way, for any real numbers a and b, we can think of the expression $|a - b|$ in geometric terms as the distance between a and b on the real number line.



This is the case $x > 0$.

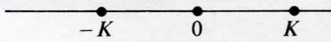


The connection between order and the modulus function is given by

Theorem 1

For all $x \in \mathbf{R}$ and all $K \geqslant 0$,

$$|x| \leqslant K \quad \text{if and only if} \quad -K \leqslant x \leqslant K.$$



Proof

As often happens with an "if and only if" assertion, we split the proof into two parts.

Suppose first that $|x| \leqslant K$. Then $0 \leqslant x \leqslant K$ if $x \geqslant 0$, while if $x < 0$ then $-x \leqslant K$ so that $-K \leqslant x < 0$. Hence in both cases $-K \leqslant x \leqslant K$.

In the opposite direction, suppose that $-K \leqslant x \leqslant K$. Then $|x| = x \leqslant K$ if $x \geqslant 0$, while if $x < 0$ then $|x| = -x$ and $x \geqslant -K$ so that $-x \leqslant K$. Hence in both cases $|x| \leqslant K$. ■

Example

Find the set S of all $x \in \mathbf{R}$ such that $|x + 3| \leqslant 4$.

Solution

We shall look at S from two different viewpoints. First of all, using Theorem 1, we see that we can interpret S as the set of $x \in \mathbf{R}$ such that

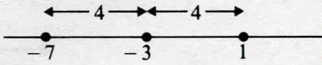
$$-4 \leqslant x + 3 \leqslant 4;$$

that is,

$$-7 \leqslant x \leqslant 1.$$

Hence, clearly, $S = [-7, 1]$.

Next, we notice that since $|x + 3| = |x - (-3)|$ we can interpret S as the set of $x \in \mathbf{R}$ whose distance from -3 is less than or equal to 4. So again we see that $S = [-7, 1]$.



It is easy to check that, for any x,y in **R**, $\left|\frac{1}{x}\right| = \frac{1}{|x|}$ and $|xy| = |x|\,|y|$. However, it is

For $\left|\frac{1}{x}\right|$ we assume $x \neq 0$.

not necessarily true that $|x + y|$ equals $|x| + |y|$ or that $|x - y|$ equals $|x| - |y|$; for example,

$$|5 + (-3)| \neq |5| + |-3| \quad \text{and} \quad |5 - (-3)| \neq |5| - |-3|.$$

It turns out, however, that what does hold for $|x + y|$ and $|x - y|$ is a key tool in many of the proofs of results in both real and complex analysis:

Theorem 2

For any $x, y \in \mathbf{R}$,

(a) $|x + y| \leqslant |x| + |y|$, and

(b) $|x - y| \geqslant ||x| - |y||$.

Remarks

Either part of Theorem 2 is often called "the Triangle Inequality" and (b) is sometimes called "the cunning form of the Triangle Inequality". Do not confuse the two different inequalities in the theorem.

You will meet the analogue for **C** of Theorem 2 in Example 6, page 42, of *Unit 1*, where the reason for the theorem's name will be explained.

The Triangle Inequality was not much used in *M203*. However, in *M332* it is vital that you are familiar with the inequality and able to use it without difficulty.

Proof of Theorem 2

(a) If $x, y \in \mathbf{R}$, then putting $K = |x|$ and $K = |y|$ in turn into the result of Theorem 1, we have that

$$-|x| \leqslant x \leqslant |x| \quad \text{and} \quad -|y| \leqslant y \leqslant |y|.$$

Adding these inequalities gives

$$-(|x| + |y|) \leqslant x + y \leqslant |x| + |y|;$$

so that, using Theorem 1 with $K = |x| + |y|$, we conclude that

$$|x + y| \leqslant |x| + |y|.$$

(b) For any $x, y \in \mathbf{R}$, it is obviously true that

$$x = (x - y) + y.$$

If we apply part (a) to this identity, it follows that

$$|x| = |(x - y) + y| \leqslant |x - y| + |y|,$$

so that

$$|x| - |y| \leqslant |x - y|. \qquad (1)$$

The same argument with x and y interchanged gives

$$|y| - |x| \leqslant |y - x|;$$

and since $|y - x| = |x - y|$, we see that

$$|y| - |x| \leqslant |x - y|. \qquad (2)$$

Thus, combining inequalities (1) and (2),

$$-|x - y| \leqslant |x| - |y| \leqslant |x - y|;$$

and so, by Theorem 1 with $K = |x - y|$,

$$||x| - |y|| \leqslant |x - y|,$$

as required. ■

Remark

Notice the above "trick" of writing x as $(x - y) + y$. It is a common trick in many proofs in analysis to regroup expressions in a more convenient way. You will see it often in this course.

Before looking at an application of the Triangle Inequality, it is convenient to introduce some more terminology. A subset S of $\mathbf{R}$ is said to be

(i) **bounded** if there exists some $K \in \mathbf{R}$ such that

$$|x| \leqslant K \qquad \text{for all } x \in S;$$

(ii) **bounded above by** M (where $M \in \mathbf{R}$) if

$$x \leqslant M \qquad \text{for all } x \in S; \text{ and}$$

(iii) **bounded below by** m (where $m \in \mathbf{R}$) if

$$m \leqslant x \qquad \text{for all } x \in S.$$

M is called an **upper bound** for S, and m a **lower bound** for S. It is not hard to show that a set S is bounded if and only if it is both bounded above and bounded below.

Example

Find an upper bound for each of the following sets.

(a) $S_1 = \{|x^2 - 4x + 1| : -1 \leqslant x \leqslant 4\}$

(b) $S_2 = \{\left|\dfrac{x}{4+3x}\right| : -1 < x \leqslant 1\}$

Solution

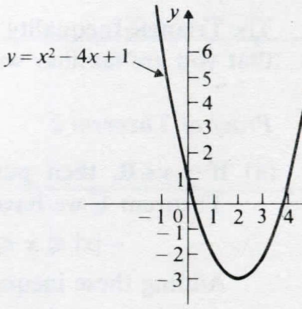
(a) Using the Triangle Inequality twice, we have that

$$\begin{aligned} |x^2 - 4x + 1| &\leqslant |x^2 + (-4x)| + 1 \\ &\leqslant |x^2| + 4|x| + 1, \end{aligned}$$

for any real x. Hence, since $|x| \leqslant 4$ whenever $x \in [-1, 4]$, it follows that, for $x \in [-1, 4]$,

$$|x^2 - 4x + 1| \leqslant 4^2 + 4 \cdot 4 + 1 = 33.$$

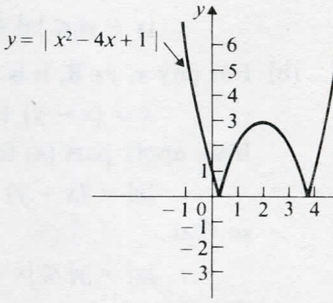
In other words, S_1 is bounded above by 33.



(b) Here we have to be more careful, and shall look at the numerator and denominator separately.

First we notice that for any $x \in (-1, 1]$,

$$|x| \leqslant 1.$$



Next we notice that we can determine an upper bound for $\dfrac{1}{|4+3x|}$ by determining a positive lower bound for $|4 + 3x|$, since, for positive M, $\dfrac{1}{|4+3x|} \leqslant M$ if and only if $|4 + 3x| \geqslant \dfrac{1}{M}$. This suggests that we might apply the "cunning form" of the Triangle Inequality to $|4 + 3x|$. Hence, by Theorem 2(b), whenever $x \in (-1, 1]$ we have

$$\begin{aligned} |4 + 3x| = |(4) - (-3x)| &\geqslant ||4| - |-3x|| \\ &= |4 - 3|x||. \qquad (1) \end{aligned}$$

We choose now to use the fact that

$$|4 - 3|x|| \geqslant 4 - 3|x|, \qquad (2)$$

rather than that $|4 - 3|x|| \geqslant 3|x| - 4$. Both inequalities are true; but, since we are looking for a positive lower bound and since $4 - 3|x| > 0$ in $(-1, 1]$, inequality (2) will be the more useful.

From (1) and (2) we deduce that

$$\begin{aligned} |4 + 3x| &\geqslant 4 - 3|x| \\ &\geqslant 4 - 3 \qquad (\text{since } x \in (-1, 1]) \\ &= 1, \end{aligned}$$

so that $\dfrac{1}{|4+3x|} \leqslant 1 \qquad$ for all $x \in (-1, 1]$.

Combining these two results, we deduce that for $x \in (-1, 1]$,

$$\left|\frac{x}{4+3x}\right| = |x| \cdot \frac{1}{|4+3x|} \leqslant 1 \cdot 1 = 1.$$

In other words, S_2 is bounded above by 1.

Remark

Since we can rewrite $x^2 - 4x + 1$ as $(x-2)^2 - 3$, it is not hard to see that 6 (for example) is also an upper bound for S_1—a smaller upper bound than the one we found. On the other hand, no number smaller than 6 will serve as an upper bound for S_1. We shall return to the question of the existence and determination of least upper bounds in Section 1.2.

Self-assessment Question

Find an upper bound for the set

$$S = \left\{\left|\frac{x^2 - x}{x^3 - 10}\right| : -2 \leqslant x < 1\right\}.$$

Solution

We shall find an upper bound for $|x^2 - x|$ and a positive lower bound for $|x^3 - 10|$; combining these will give us an upper bound for S.

First, whenever $x \in [-2, 1)$, we have that

$$\begin{aligned} |x^2 - x| &\leqslant |x|^2 + |-x| && \text{(by the Triangle Inequality)} \\ &\leqslant 2^2 + 2 && \text{(as } |x| \leqslant 2\text{)} \\ &= 6. \end{aligned}$$

Next we use the "cunning form" of the Triangle Inequality to deduce that, for $x \in [-2, 1)$,

$$\begin{aligned} |x^3 - 10| &\geqslant \big||x^3| - |10|\big| \\ &\geqslant 10 - |x|^3 && \text{(by Theorem 1)} \\ &\geqslant 10 - 8 \\ &= 2. \end{aligned}$$

We choose $10 - |x|^3$ rather than $|x|^3 - 10$, since $10 - |x|^3$ is positive in $[-2, 1)$.

Hence, for $x \in [-2, 1)$, $\dfrac{1}{|x^3 - 10|} \leqslant \dfrac{1}{2}$.

Consequently,

$$\left|\frac{x^2 - x}{x^3 - 10}\right| \leqslant \frac{6}{2} = 3,$$

so that 3 is an upper bound for S.

1.2 The Least Upper Bound Property of R

We noticed in the last example that if a subset S of **R** is bounded above by some real number M, the value of M is not unique. For example, $M + 1$ and $M + \pi$ are also upper bounds for S. Often however, given a set S, what we need is not simply *any* upper bound, but the *least* upper bound that we can find (if indeed such a thing exists).

Definition

Given a non-empty subset S of **R** which is bounded above, we call u a **least upper bound** or **supremum** for S (denoted by the symbol l.u.b. S or sup S) if

(a) u is an upper bound for S; and

(b) $u \leqslant M$ for any upper bound M for S.

We can put this definition in the following equivalent form:

(a′) $s \leqslant u$ for every $s \in S$; and

(b′) if $u' < u$, then there is at least one $s \in S$ with $s > u'$.

In other words, (a′) ensures that u is an upper bound for S and (b′) ensures that no smaller number can be an upper bound for S. Notice too that every real number larger than u is also an upper bound for S.

Examples

(a) The set **R** has no upper bound (and no lower bound), and so no least upper bound.

(b) The set $S = (-\infty, 3)$ has no lower bound, but any real number $M \geqslant 3$ is an upper bound for S. In fact, sup $S = 3$. For, since $s \leqslant 3$ for every $s \in S$, (a′) is satisfied. Next, for any real number $u' < 3$, the number $s = \frac{1}{2}(u' + 3)$ belongs to S and is strictly greater than u'; and so (b′) is also satisfied.

(c) If $c < d$, the least upper bound of both the intervals (c, d) and $[c, d]$ is d.

In view of (b) and (c), we notice that the least upper bound of a set S in **R** may or may not belong to S.

The real numbers have a property over and above their arithmetic and order-related properties (all of which are shared by the rational numbers) which distinguishes **R** from **Q**, namely the Least Upper Bound Property of **R**. We may express this intuitively as "there are no gaps in the real line"—unlike the gaps in **Q** caused by the existence of irrational numbers such as $\sqrt{2}$.

Q represents the set of rationa numbers.

Least Upper Bound Property of R

Any non-empty subset of **R** which is bounded above has a least upper bound.

This property is really rather subtle. However, it may help to think of it in the following way.

Given any finite collection $S = \{x_1, \ldots, x_n\}$ of real numbers, the collection always contains one element which is greater than or equal to all the other elements; that is, a maximum element denoted by

$$\max S, \text{ or } \max(x_1, \ldots, x_n).$$

An alternative notation is $\max\{x_1, \ldots, x_n\}$.

However, if we are dealing with a subset S of **R** which contains *infinitely* many elements, such as the intervals $(0, 1)$ and $[0, 1]$, S *may or may not* possess a maximum element. Hence when we deal with infinite sets $S \subseteq \mathbf{R}$ we often talk about the least upper bound of S rather than the maximum element of S. The Least Upper Bound Property of **R** says that for any non-empty set S that is bounded above, sup S always exists.

We remark, without proof, that if a set possesses a least upper bound, it is quite easy to verify (from the definition of least upper bound) that it is necessarily unique.

Theorem 3

Let S be a non-empty subset of $\mathbf{R}$ that is bounded above. Then there exists a unique real number u such that $u = \sup S$.

The properties of sets bounded above have direct analogues for sets that are bounded below; we look next at these.

Definition

Given a non-empty subset S of $\mathbf{R}$ which is bounded below, we call v a **greatest lower bound** or **infimum** for S (denoted by the symbol g.l.b. S or $\inf S$) if

(a) v is a lower bound for S; and

(b) $v \geqslant m$ for any lower bound m for S.

As before, we can reformulate this definition in an equivalent form:

(a′) $s \geqslant v$ for every $s \in S$; and

(b′) if $v' > v$, then there is at least one $s \in S$ with $s < v'$.

So (a′) ensures that v is a lower bound for S and (b′) ensures that no larger number can be a lower bound for S. Also, every real number less than v is a lower bound for S.

The following theorem is an immediate consequence of the Least Upper Bound Property of $\mathbf{R}$; it is sometimes called the Greatest Lower Bound Property of $\mathbf{R}$.

Theorem 4

Let S be a non-empty subset of $\mathbf{R}$ that is bounded below. Then S has a greatest lower bound.

Remark

The strategy of the proof is to use the fact that m is a lower bound for S if and only if $-m$ is an upper bound for some other set, to which we shall apply the Least Upper Bound Property of $\mathbf{R}$.

Proof of Theorem 4

Let $T = \{x \in \mathbf{R}\colon -x \in S\}$.

If m is a lower bound for S, then $m \leqslant s$ for all $s \in S$; so that if $x \in T$ then $m \leqslant -x$, and hence $x \leqslant -m$. Thus T is bounded above and $-m$ is an upper bound for T.

It follows from the Least Upper Bound Property of $\mathbf{R}$ that T has a least upper bound, u say; so that $x \leqslant u \leqslant -m$ for each $x \in T$. Hence $-u \leqslant -x$, so that $-u$ is a lower bound for S.

But if m is any lower bound for S then $-m$ is an upper bound for T, so that (by definition of least upper bound) $u \leqslant -m$, and hence $m \leqslant -u$. Hence $-u$ is the greatest lower bound for S. ■

Example

For the set $S = (2, \infty)$, show that $\inf S = 2$.

Solution

For each element s of S, $s \geqslant 2$. (In fact, $s > 2$, but this does not matter!) So (a′) is satisfied by 2.

Next, for any real number $v' > 2$, the number $\frac{1}{2}(v' + 2)$ belongs to S and is strictly less than v'; and so (b′) is also satisfied by 2. It follows that $\inf S = 2$.

For the proof of our next Example, we introduce another piece of terminology.

Given any finite collection $\{x_1, \ldots, x_n\}$ of real numbers, we denote its **minimum element** (that is, the element which is less than or equal to all the other elements) by the symbol $\min(x_1, \ldots, x_n)$.

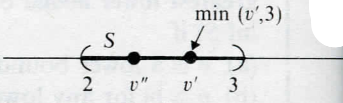
Example

For the set $S = (2,3)$, show that $\inf S = 2$.

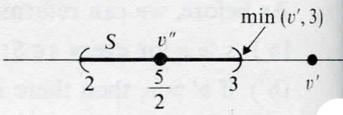
Solution

As in the previous Example, 2 is certainly a lower bound for S. Now let v' be any real number with $v' > 2$.

Then the real number $v'' = \frac{1}{2}(2 + \min(v', 3))$ belongs to S (as illustrated in the marginal figures — the details can be written out carefully without much trouble), and v'' is strictly greater than 2; and so (b$'$) is also satisfied.



Case 1: $2 < v' < 3$



Case 2: $3 \leqslant v'$

It follows that $\inf S = 2$.

Remark

For the set $S = [2, 3)$, $\inf S = 2$. This is easy to see since 2 is a lower bound for S and no larger number can be a lower bound (for it would be larger than the element 2 of S).

Comparing this result with that of the previous Example, we see that the greatest lower bound of a set may or may not belong to the set itself.

The terms least upper bound and greatest lower bound of a set were introduced in *M203, Analysis Block A, Unit 1, Numbers,* Section 4.2.

Just as the notions of supremum and infimum of a set S generalize those of maximum and minimum elements of S, we can talk about the supremum and infimum of a function f defined on a set $S \subseteq \mathbf{R}$ as generalizations of the maximum and minimum values of f on S.

Definition

Let f be a function with domain $S \subseteq \mathbf{R}$ and codomain $\mathbf{R}$, and denote its image $f(S)$ by T. Then the **least upper bound** or **supremum** of f on S is defined to be

$$\sup_{x \in S} f = \sup T,$$

and the **greatest lower bound** or **infimum** of f on S is defined to be

$$\inf_{x \in S} f = \inf T.$$

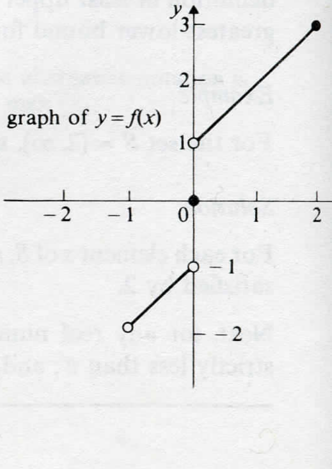
In the same way, we say that the function f is **bounded, bounded above** by M or **bounded below** by m on the set S if T is bounded, bounded above by M or bounded below by m, respectively.

Example

Let $f\colon (-1, 2] \longrightarrow \mathbf{R}$ be defined by

$$f\colon x \longrightarrow \begin{cases} x + 1, & \text{if } 0 < x \leqslant 2, \\ 0, & \text{if } x = 0, \\ x - 1, & \text{if } -1 < x < 0. \end{cases}$$

What are $\inf\limits_{x \in (-1,2]} f$ and $\sup\limits_{x \in (-1,2]} f$?



Outline Solution

In this case,

$$T = f((-1, 2]) = (-2, -1) \cup \{0\} \cup (1, 3];$$

and it is then quite easy to prove carefully that

$$\inf T = -2 \quad \text{and} \quad \sup T = 3,$$

so that

$$\inf_{x \in (-1,2]} f = -2 \quad \text{and} \quad \sup_{x \in (-1,2]} f = 3.$$

Remarks

(i) Notice that f actually attains its supremum 3 at the point $x = 2$, but does not attain its infimum (that is, there is no $x \in (-1, 2]$ for which $f(x) = -2$).

(ii) In almost all our applications of supremum and infimum in *M332*, the functions f that we meet will be both continuous and differentiable.

We claim that if f: $[a, b] \longrightarrow \mathbf{R}$ is continuous on a closed interval $[a, b] \subset \mathbf{R}$, then

(i) $f([a, b])$ is bounded; and

(ii) there exist points c, d in $[a, b]$ such that

$$f(c) = \sup_{x \in [a,b]} f \quad \text{and} \quad f(d) = \inf_{x \in [a,b]} f.$$

This is Theorem 18 on page 47.

In other words, in this special case the ideas of maximum and supremum and of minimum and infimum coincide.

Of course, if f is also differentiable, one of the best ways of finding c and d, and hence $\sup f$ and $\inf f$, is by determining where the derivative f' vanishes and comparing the values of f at these points with its values at the end-points a and b.

Example

Let the function f: $[-1, 1] \longrightarrow \mathbf{R}$ be defined by

$$f: x \longrightarrow x(1 - x).$$

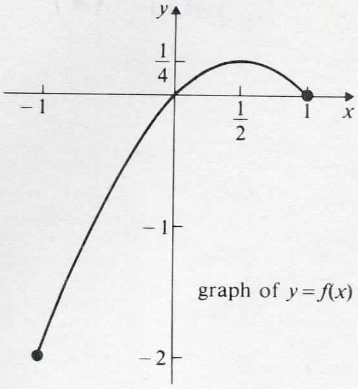
Determine $\sup_{x \in [-1,1]} f$ and $\inf_{x \in [-1,1]} f$.

Solution

f is differentiable on $(-1, 1)$, and $f'(x) = 1 - 2x$ which vanishes when $x = \frac{1}{2}$. So, following the usual methods, it follows that $\sup_{x \in [-1,1]} f$ and $\inf_{x \in [-1,1]} f$ are two of the values in the list

$$f(-1),\ f(1) \text{ and } f(\tfrac{1}{2});$$

that is, -2, 0 and $\frac{1}{4}$.



graph of $y = f(x)$

Consequently,

$$\sup_{x \in [-1,1]} f = \tfrac{1}{4} = f(\tfrac{1}{2})$$

and

$$\inf_{x \in [-1,1]} f = -2 = f(-1).$$

i.e. we look at the end-points plus points where $f'(x) = 0$.

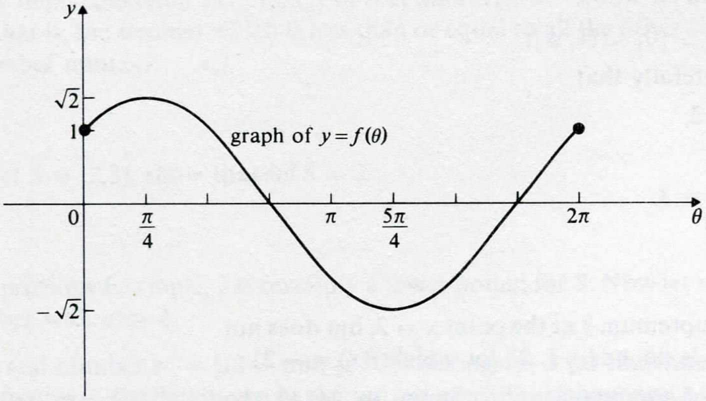
Self-assessment Question

Let the function f: $[0, 2\pi] \longrightarrow \mathbf{R}$ be defined by

$$f: \theta \longrightarrow \cos\theta + \sin\theta.$$

Determine $\sup_{x \in [0,2\pi]} f$ and $\inf_{x \in [0,2\pi]} f$, and state whether or not each is attained.

Solution



f is differentiable on $(0, 2\pi)$, and $f'(\theta) = -\sin\theta + \cos\theta$ which vanishes whenever $\tan\theta = 1$; in other words, when $\theta = \frac{\pi}{4}, \frac{5\pi}{4}$. Hence $\sup_{x\in[0,2\pi]} f$ and $\inf_{x\in[0,2\pi]} f$ are two of the following values of f:

$$f(0), f(2\pi), f\left(\frac{\pi}{4}\right) \text{ and } f\left(\frac{5\pi}{4}\right);$$

that is,

$$1, 1, \sqrt{2} \text{ and } -\sqrt{2}.$$

Hence

$$\sup_{[0,2\pi]} f = f\left(\frac{\pi}{4}\right) = \sqrt{2}$$

and

$$\inf_{[0,2\pi]} f = f\left(\frac{5\pi}{4}\right) = -\sqrt{2}.$$

1.3 The Archimedean Property of R

In this section we look in a slightly different way at the fact that "the real line contains no gaps", which we discussed in the previous section; all our results will therefore be consequences of the Least Upper Bound Property of **R**.

Theorem 5 (The Archimedean Property of the real numbers)

The set **N** of natural numbers is not bounded above.

$\mathbf{N} = \{0, 1, 2, 3, \ldots\}$.

Proof

If **N** were bounded above, there would exist some real number $u = \sup \mathbf{N}$.

Now, for each $n \in \mathbf{N}$, $n + 1$ is also in **N** so that $n + 1 \leqslant u$. In other words, for each $n \in \mathbf{N}$, $n \leqslant u - 1$ so that $u - 1$ is also an upper bound for **N**. This contradicts part (b) of the definition of $\sup \mathbf{N}$.

It follows that **N** is not bounded above. ■

Our main uses of the Archimedean Property of **R** will be based on the following result:

Theorem 6

Let ε be any given positive real number. Then there exists some $n \in \mathbf{N}$ such that

$$0 < \frac{1}{n} < \varepsilon.$$

Proof

For every $n \in \mathbf{N}$, it is true that $0 < \frac{1}{n}$, so we have only to prove the right-hand inequality.

Now $\frac{1}{\varepsilon}$ is a positive real number. Since, by Theorem 5, **N** is not bounded above, there must exist some integer $n \in \mathbf{N}$ such that $n > \frac{1}{\varepsilon}$. It follows that $\frac{1}{n} < \varepsilon$, as required. ■

Remark

The result of Theorem 6 is most useful when ε is small; it guarantees that however small ε is (provided it is positive), there is some n with $\frac{1}{n}$ smaller than ε.

Example

For the set $S = \left\{\frac{1}{n} : n \in \mathbf{N}\right\}$, find $\sup S$ and $\inf S$.

Solution

$S = \{1, \frac{1}{2}, \frac{1}{3}, \ldots\}$, so it looks as though the least upper bound for S is 1, since $1 \in S$ and $\frac{1}{n}$ gets smaller as n gets larger. Moreover, since all the elements of S are non-negative and $\frac{1}{n}$ "approaches" 0 as n gets larger, it looks likely that the greatest lower bound for S is 0. We now verify that both guesses are correct!

First, $\frac{1}{n} \leqslant 1$ for all $n \in \mathbf{N}$: so that 1 is certainly an upper bound for S. But since 1 is itself an element of S, no number less than 1 will serve as an upper bound for S, so that 1 must be the least upper bound for S. That is, $\sup S = 1$.

Next, $\dfrac{1}{n} \geqslant 0$ for all $n \in \mathbf{N}$; so that 0 is certainly a lower bound for S. Theorem 6 then shows that no positive number, ε say, can be a lower bound for S, as there is an n with $\dfrac{1}{n} < \varepsilon$. Therefore 0 must be the greatest lower bound for S. That is, $\inf S = 0$.

Remark

In this case, $\sup S$ lies in S, but $\inf S$ does not.

We end this section with a result which might seem so trivial that it is not worth stating. In fact, we shall find that it is used regularly later in the course, both for $\mathbf{R}$ (as we discuss it below) and for $\mathbf{C}$. (For example, it is the final ingredient in the proof of Cauchy's Formula on page 75 of *Unit 5*).

Theorem 7

Let x be a real number (independent of ε) with the property that $|x| < \varepsilon$ for every positive number ε. Then $x = 0$.

Proof

We shall use a contradiction argument. So, suppose that $x \neq 0$. Then $|x| > 0$ by definition of the modulus function; so that by Theorem 6, there exists some $n_0 \in \mathbf{N}$ with

$$0 < \frac{1}{n_0} < |x|.$$

Hence, for every $\varepsilon > 0$,

$$\frac{1}{n_0} < |x| < \varepsilon;$$

in other words, there exists some integer $n_0 \in \mathbf{N}$ such that, for every choice of $\varepsilon > 0$,

$$\frac{1}{n_0} < \varepsilon. \qquad (*)$$

In particular, $(*)$ must hold for the choice $\varepsilon = \dfrac{1}{n_0}$, which it clearly does not.

This contradiction shows that we must have $x = 0$. ■

2 SEQUENCES

A sequence $\{s_n\}$ of real numbers is an infinite list of real numbers $s_1, s_2, s_3, \ldots$. For example, the following are all sequences of real numbers:

(i) $1, \frac{1}{2}, \frac{1}{3}, \ldots, \frac{1}{n}, \ldots;$

(ii) $1, 0, 1, 0, 1, 0, \ldots;$

(iii) $1, 2, 3, 4, 5, \ldots, n, \ldots$.

Of course, as we do not wish to write down all the infinitely many elements of a sequence, we use convenient short-hand devices such as continuing a list of the first few elements of the sequence by three dots when it is obvious how the list should be continued, or by giving a simple formula for the nth element in the list. For example, the nth element in each of the above sequences $\{s_n\}$ can be specified as follows:

(i) $s_n = \frac{1}{n}$; (ii) $s_n = \frac{1}{2}\{1 + (-1)^{n+1}\}$; (iii) $s_n = n$.

Sequences arise frequently in numerical mathematics; for example, s_n may be the answer at the nth stage in some method of finding a root of an algebraic equation by successive approximation. They are just as important in pure mathematics, as you will see. In most situations where sequences are used, the interesting thing is not the value of any particular element of the sequence, but rather the existence and value of any real number to which the sequence *converges* (or *tends*, or *approaches*).

2.1 Convergence of Sequences

We now look at the definition of the convergence of a given sequence of real numbers in terms of ε and N, and then at the roles of ε and N in this definition.

Definition

The sequence $\{s_n\}$ of real numbers **converges** to the real number l if, given any real number $\varepsilon > 0$, there exists a natural number N such that

$$|s_n - l| < \varepsilon \quad \text{for all } n \geqslant N. \qquad (1)$$

Sometimes we shall write "$\lim_{n\to\infty} s_n = l$" or "$s_n \to l$ as $n \to \infty$" instead of "$\{s_n\}$ converges to l". We call l the **limit** of the sequence $\{s_n\}$.

Notice two things about the ε and N that appear in the definition. Firstly, the order in which ε and N occur is crucial: the definition requires that some $\varepsilon > 0$ must first be specified, and that *then* there must be a natural number N such that the condition (1) holds. (For different given values of ε, the corresponding values of N that we choose may well be different.) Secondly, to show that a sequence $\{s_n\}$ converges to l, it is not enough to show that for any given (pre-assigned) value of $\varepsilon > 0$ we can find an N such that

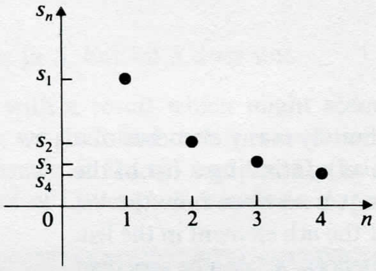
$$|s_n - l| < \varepsilon \qquad \text{for some } n \geqslant N;$$

for (1) must hold for *all* $n \geqslant N$.

Example

Show that the sequence $\{s_n\}$, where $s_n = \frac{1}{n}$, converges to the limit 0.

Solution



Let us first of all look at the definition (1) of convergence in numerical terms. When $s_n = \frac{1}{n}$ and $l = 0$, the sequence $\{s_n\}$ will converge to 0 if, for any given $\varepsilon > 0$, there exists a natural number N such that

$$|s_n - 0| < \varepsilon \qquad \text{for all } n \geqslant N;$$

that is,

$$\left|\frac{1}{n} - 0\right| < \varepsilon \qquad \text{for all } n \geqslant N. \qquad (2)$$

If we were told initially that $\varepsilon = \frac{1}{10}$, then we would want to find an N such that $\frac{1}{n} < \frac{1}{10}$ for all $n \geqslant N$; so, obviously, we could satisfy (2) by choosing $N = 11$.

Of course, any natural number N larger than 10 would have done equally well for our purpose; it just happened that the number 11 sprang to mind and that (2) was indeed satisfied with the choice $N = 11$.

If we were told to satisfy (2) in the special case that $\varepsilon = 10^{-6}$, in exactly the same way we see that (2) will be satisfied for all $n \geqslant N$ for any choice of natural number N larger than 10^6; for example, we could choose $N = 10^6 + 1$ or even $N = 10^{10}$.

However, in this way we are not yet *proving* that $\{s_n\}$ converges to 0. We have to prove that, for *any* given value of $\varepsilon > 0$, an N exists for which (2) holds for all $n \geqslant N$.

The condition that $\left|\frac{1}{n} - 0\right| < \varepsilon$ can be rewritten as $\frac{1}{n} < \varepsilon$ or $n > \frac{1}{\varepsilon}$; and this suggests that if we choose N larger than $\frac{1}{\varepsilon}$ then (2) will necessarily hold for all $n \geqslant N$.

So, let $\varepsilon > 0$ be given. Then choose N to be the smallest natural number greater than $\frac{1}{\varepsilon}$. It follows that, for all $n \geqslant N$,

$$\begin{aligned} |s_n - 0| &= \frac{1}{n} \\ &\leqslant \frac{1}{N} \qquad (\text{as } n \geqslant N) \\ &< \varepsilon \qquad \left(\text{as } N > \frac{1}{\varepsilon}\right); \end{aligned}$$

so that $\{s_n\}$ does, indeed, converge to 0.

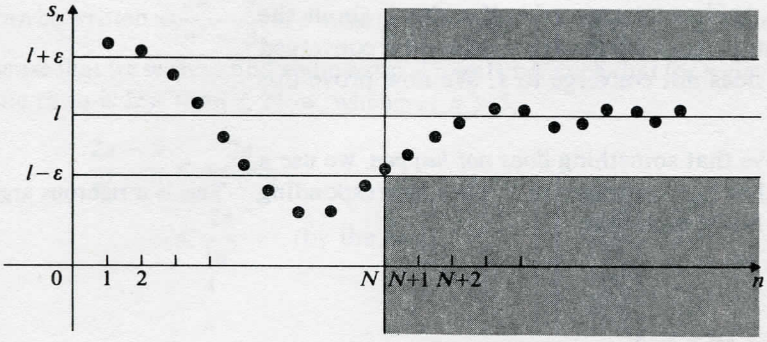
We may think of the definition of limit geometrically. If we plot the values of s_n against n for a sequence which converges to some number l, we know that for any pre-assigned value of ε, s_n must *eventually* satisfy the condition $|s_n - l| < \varepsilon$; that is

$$-\varepsilon < s_n - l < \varepsilon,$$

or

$$l - \varepsilon < s_n < l + \varepsilon.$$

In other words, from some index N onwards, the corresponding values of the elements of the sequence lie within a (narrow) horizontal strip of width 2ε centred at the value l.



If we choose the strip to be very narrow, the minimum value of N that will work is likely to be quite large. Of course, for any given value of ε, there is no need for us to find the *minimum* value of N for which (1) holds, and usually we shall not bother to do so. It is also clear from the diagram that the convergence or otherwise of a sequence is unchanged if a *finite* number of terms in the sequence are added, deleted or changed—all that happens is that the value of N in (1) may be changed.

Example

Show that the sequence $\{s_n\}$, where $s_n = \dfrac{n^2}{n^2+1}$, converges.

Solution

If we rewrite the formula for s_n as $s_n = 1 - \dfrac{1}{n^2+1}$, it is easy to guess what must happen; as n gets larger, $1/(n^2+1)$ "approaches" 0 and so s_n "approaches" the value 1. As it is possible for our intuition to be mislead sometimes, we check that this guess is correct using the Definition.

This is motivation.

Let $\varepsilon > 0$ be given. The Definition tells us to look at the difference $|s_n - 1|$; here $|s_n - 1| = \dfrac{1}{n^2+1}$. How can we find N such that

This is a rigorous argument.

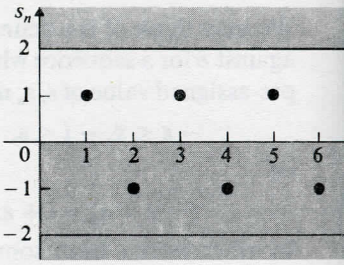
$$\frac{1}{n^2+1} < \varepsilon \qquad \text{for all } n \geqslant N?$$

Notice first that $\dfrac{1}{n^2+1} < \dfrac{1}{n^2}$, so that $\dfrac{1}{n^2+1} < \varepsilon$ whenever $\dfrac{1}{n^2} < \varepsilon$. Now $\dfrac{1}{n^2} < \varepsilon$ will be the case whenever $n^2 > \dfrac{1}{\varepsilon}$ or $n > \dfrac{1}{\varepsilon^{1/2}}$. So, if we choose N to be the least natural number greater than $\varepsilon^{-1/2}$, it follows that, for all $n \geqslant N$,

$$\begin{aligned} |s_n - 1| &= \frac{1}{n^2+1} \\ &< \frac{1}{n^2} \leqslant \frac{1}{N^2} < \varepsilon. \end{aligned}$$

It now follows from the Definition that $\{s_n\}$ converges to 1.

Just as we can use the Definition to prove that a sequence $\{s_n\}$ converges we can also use it to prove that a sequence does not converge.



Example

Show that the sequence $\{s_n\}$, where $s_n = (-1)^{n+1}$, that is the sequence $\{1, -1, 1, -1, \ldots\}$, does not converge to $+1$.

Solution

This is motivation.

If $\{s_n\}$ converged to 1 as $n \to \infty$, we should expect that "eventually" all values of s_n would be "near to" 1. This is not the case, as half of them are distance 2 from $+1$, and, in particular, do not lie in a horizontal strip of half width 1 about the value 1. Since s_n should eventually stay out of the shaded territory if it converged to 1, we see "intuitively" that $\{s_n\}$ does not converge to 1. We now prove this carefully, using the Definition.

This is a rigorous argument.

As we often do when asked to prove that something *does not happen*, we use a contradiction argument. Suppose that $\{s_n\}$ converges to 1. Then, corresponding to any given $\varepsilon > 0$ there exists an $N \in \mathbf{N}$ such that

$$|s_n - 1| < \varepsilon \qquad \text{for all } n \geqslant N;$$

that is,

$$|(-1)^{n+1} - 1| < \varepsilon \qquad \text{for all } n \geqslant N.$$

This has to hold for both even and odd $n \geqslant N$; the inequality then reads $2 < \varepsilon$, $0 < \varepsilon$ respectively. This suggests that for a contradiction we might agree to choose $\varepsilon = 1$ initially (or indeed any ε with $\varepsilon \leqslant 2$). Our hypothesis then implies the existence of some natural number N such that

$$2 < 1 \text{ for even } n \geqslant N, \text{ and } 0 < 1 \text{ for odd } n \geqslant N.$$

The first inequality is absurd! So $\{s_n\}$ cannot converge to 1 after all.

Remark

This is the strategy for showing that a sequence $\{s_n\}$ does not converge to a given value l. What we have to do is to find a value of ε such that the condition

$$|s_n - l| < \varepsilon \qquad \text{for all } n \geqslant N$$

cannot be satisfied for any choice whatever of natural number N.

Self-assessment Question

(a) Find a value of l such that the sequence $\{s_n\}$, where $s_n = \dfrac{3n^2 - 2n}{n^2 - 3}$, converges to l; and prove that $\{s_n\}$ does indeed converge to l.

(Hint: You may find the following inequality useful: $n^2 - 3 > \frac{1}{2}n^2$ for all $n > 2$.)

(b) Prove that the sequence $\{t_n\}$, where $t_n = n$, does not converge to any real number l.

Solution

Again, this is motivation.

(a) If a sequence $\{s_n\}$ is to be convergent, we need only consider the values of s_n for large indices n. Now, when n is large, $3n^2 - 2n$ is more or less the same as $3n^2$, and $n^2 - 3$ is more or less the same as n^2; hence s_n is more or less the same as $\dfrac{3n^2}{n^2} = 3$. We can make the phrase "more or less" a bit more precise as follows.

Write s_n in the form $s_n = \dfrac{3 - \dfrac{2}{n}}{1 - \dfrac{3}{n^2}}$. As n gets larger and larger the numerator gets closer and closer to the value 3, and similarly the denominator to the value 1; hence s_n approaches the value 3. Hence it is reasonable to guess that the sequence converges to the value $l = 3$. We next prove this carefully.

We know from the definition of convergence that we should look at the difference

$$|s_n - 3| = \left|\frac{3n^2 - 2n}{n^2 - 3} - 3\right| = \frac{|9 - 2n|}{|n^2 - 3|}.$$

We are interested in this difference only when n is large. So, if we agree that we shall consider only those indices n greater than or equal to 5, this difference can be written as $\dfrac{2n - 9}{n^2 - 3}$. We want to find an upper bound for this ratio, in the sense that we wish to find some natural number N such that for a given $\varepsilon > 0$ the ratio is less than ε. Now, whenever $n \geqslant 5$,

$$\begin{aligned} \frac{2n - 9}{n^2 - 3} &< \frac{2n}{n^2 - 3} \\ &\leqslant \frac{2n}{\frac{1}{2}n^2} \qquad \text{(by the hint)} \\ &= \frac{4}{n}; \end{aligned}$$

this will be less than ε if, in addition, we require that $n > \dfrac{4}{\varepsilon}$. This shows us that, whatever value ε takes, we wish to ensure that we deal only with n for which both $n \geqslant 5$ and $n > \dfrac{4}{\varepsilon}$.

So, let $\varepsilon > 0$ be given. Choose N to be the least natural number greater than $\max\left(5, \dfrac{4}{\varepsilon}\right)$ (so that all the above inequalities hold). Then for all $n \geqslant N$,

This is the rigorous argument.

$$\begin{aligned} |s_n - 3| = \frac{|9 - 2n|}{|n^2 - 3|} &< \frac{4}{n} \qquad \text{(since } n > 5\text{)} \\ &< \varepsilon \qquad \left(\text{since } n > \frac{4}{\varepsilon}\right); \end{aligned}$$

from which it follows that $\{s_n\}$ converges to 3, as desired.

(b) We use a contradiction argument.

Suppose there is some real number l such that $\{t_n\}$ converges to l. Then, for any given $\varepsilon > 0$, there exists a corresponding natural number N such that

$$|t_n - l| < \varepsilon \qquad \text{for all } n \geqslant N.$$

In particular, if we take $\varepsilon = 1$ (any specific ε will do!), there exists some N such that

$$|n - l| < 1 \qquad \text{for all } n \geqslant N,$$

and so

$$n < l + 1 \qquad \text{for all } n \geqslant N.$$

It follows, then, that since n is bounded above by $l + 1$ for all $n \geqslant N$, and by N for all $n < N$,

$$n < \max(l + 1, N) \qquad \text{for all } n \in \mathbf{N}.$$

This contradiction with the Archimedean Property of $\mathbf{R}$ shows that $\{t_n\}$ cannot converge.

Part (b) of this last SAQ raises the interesting question of whether a sequence is convergent if and only if it is bounded.

Self-assessment Question

Show that the sequence $\{s_n\}$, where $s_n = (-1)^{n+1}$, does not converge to any finite number l.

Solution

Following the advice in the Example on page 22, we use a contradiction argument.

Suppose that a real number $l = \lim_{n\to\infty} s_n$ exists. Then, corresponding to *any* given $\varepsilon > 0$, there exists an $N \in \mathbf{N}$ such that

$$|s_n - l| < \varepsilon \qquad \text{for all } n \geqslant N;$$

that is,

$$|(-1)^{n+1} - l| < \varepsilon \qquad \text{for all } n \geqslant N.$$

Whatever N is, there are both even and odd n that must satisfy this inequality ($N + 1$ and $N + 2$, for instance); hence we must have

$$|1 + l| < \varepsilon \quad \text{and} \quad |1 - l| < \varepsilon \qquad (\text{since } |-1 - l| = |1 + l|).$$

n has now disappeared; and whatever l is, both inequalities have to be satisfied for every $\varepsilon > 0$.

Whenever you see expressions like $1 + l$, $1 - l$, you should immediately think of applying the Triangle Inequality to one of the expressions $(1 + l) + (1 - l)$ or $(1 + l) - (1 - l)$. Here we use the first: so that, for all $\varepsilon > 0$,

$$\begin{aligned} 2 &= |(1 + l) + (1 - l)| \\ &\leqslant |1 + l| + |1 - l| < 2\varepsilon. \end{aligned}$$

Consequently, if at the beginning we choose $\varepsilon = \frac{1}{2}$, we get the desired contradiction.

Although boundedness does not, in itself, imply convergence, for special types of sequences, *monotonic* sequences, it is actually true that if a sequence is bounded then it is convergent.

Definition

A sequence $\{s_n\}$ is **increasing** if $s_{n+1} \geqslant s_n$ for all $n \in \mathbf{N}$, and **decreasing** if $s_{n+1} \leqslant s_n$ for all $n \in \mathbf{N}$. It is called **monotonic** if it has either of these properties.

Theorem 8

Let $\{s_n\}$ be an increasing (decreasing) sequence that is bounded above (below). Then $\{s_n\}$ is convergent.

Proof

We give the details for an increasing sequence only; for a decreasing sequence the inequalities are reversed and upper bounds replaced by lower bounds, but otherwise the argument is the same.

Let $\{s_n\}$ be an increasing sequence that is bounded above. By the Least Upper Bound Property of $\mathbf{R}$, the set S of elements of the sequence possesses a supremum, l say. We now check that the sequence actually converges to l.

Choose any $\varepsilon > 0$. Since $l - \varepsilon < l$, it follows that $l - \varepsilon$ cannot be an upper bound for S since l is the *least* upper bound for S. Hence there exists some index N for which $l - \varepsilon < s_N$. Consequently, as $\{s_n\}$ is an increasing sequence,

$$l - \varepsilon < s_n \qquad \text{for all } n \geqslant N;$$

and, since l is an upper bound for S, we also have that

$$\begin{aligned} s_n &\leqslant l \qquad \text{for all } n \\ &< l + \varepsilon. \end{aligned}$$

Combining the last two inequalities, we see that, for all $n \geqslant N$, $l - \varepsilon < s_n < l + \varepsilon$; in other words,

$$|s_n - l| < \varepsilon \qquad \text{for all } n \geqslant N.$$

It follows that $\{s_n\}$ converges to l, as desired. ■

Although the boundedness of a sequence $\{s_n\}$ does not imply its convergence, the converse result holds.

Theorem 9

Let the sequence $\{s_n\}$ converge to l. Then the sequence is bounded.

Proof

Any choice of ε will do, as you will see when we choose K below.

Since $\{s_n\}$ converges to l, corresponding to any given $\varepsilon > 0$ there exists a natural number N such that $|s_n - l| < \varepsilon$ for all $n \geqslant N$. In particular, if we choose $\varepsilon = 1$, then there exists a corresponding value of N such that

$$|s_n - l| < 1 \qquad \text{for all } n \geqslant N.$$

It follows that, for all $n \geqslant N$,

Remember that on page 9 we said this was a useful trick!

$$\begin{aligned} |s_n| &= |(s_n - l) + l| \\ &\leqslant |s_n - l| + |l| \qquad \text{(by the Triangle Inequality)} \\ &< 1 + |l|. \end{aligned}$$

In addition, it is obvious that, for all $n < N$,

$$\begin{aligned} |s_n| &\leqslant \max\,(|s_1|, |s_2|, \ldots, |s_{N-1}|) \\ &= M, \quad \text{say.} \end{aligned}$$

Consequently, if we put $K = \max\,(1 + |l|, M)$ we conclude that

$$|s_n| \leqslant K \qquad \text{for all } n \in \mathbf{N},$$

so that the sequence $\{s_n\}$ is bounded. ■

Remark

Had we chosen a different value of ε, the value of K would have been different, but otherwise the proof is the same.

Self-assessment Question

Prove that the sequence $\{s_n\}$, where $s_n = \dfrac{1}{2} + \dfrac{1}{(2!)^2 2^2} + \dfrac{1}{(3!)^3 2^3} + \ldots + \dfrac{1}{(n!)^n 2^n}$, converges. (Do not attempt to find the value of $\lim\limits_{n \to \infty} s_n$.)

Solution

Since, for each n,

$$s_{n+1} - s_n = \frac{1}{((n+1)!)^{n+1} 2^{n+1}} > 0,$$

the sequence is increasing; and since, for each n,

$$\begin{aligned} s_n &\leqslant \frac{1}{2} + \frac{1}{2^2} + \ldots + \frac{1}{2^n} \\ &< \sum_{n=1}^{\infty} \frac{1}{2^n} = 1, \end{aligned}$$

the sequence $\{s_n\}$ is bounded above (by 1).

It follows from Theorem 8 that the sequence must be convergent.

Remark

For future reference, we mention here the following useful results;

(i) For any real number x, $\lim_{n \to \infty} \frac{x^{n+1}}{(n+1)!} = 0$.

We use this on page 39, for exam

(ii) For any $a > 0$, $\lim_{n \to \infty} a^{1/n} = 1$.

2.2 Some Properties of Convergent Sequences

In this section, we check that combinations of sequences such as sums and products have the expected convergence properties. Our main result is:

Theorem 10

Let $\{s_n\}$, $\{t_n\}$ be sequences converging to l_1, l_2 respectively, and let $k \in \mathbf{R}$. Then, as $n \to \infty$,

(i) $ks_n \to kl_1$;

(ii) $s_n + t_n \to l_1 + l_2$; and

(iii) $s_n t_n \to l_1 l_2$.

Before looking at the proofs of (i)–(iii), we make an observation about the "ε" which appears in the inequality (1) in the definition of convergence of a sequence. This trivial-looking result will save us from a great deal of complication in such proofs.

Lemma 11 (First version of the "$K\varepsilon$ Lemma")

Suppose that, corresponding to each given $\varepsilon > 0$, there exists a natural number N such that

$$|s_n - l| < K\varepsilon \qquad \text{for all } n \geqslant N,$$

where K is a (positive) real number independent of ε and N. Then the sequence $\{s_n\}$ converges to the limit l.

Remark

In applications, we shall often find that K is a real number (such as 2 or 25), or a function depending on l (such as $1 + |l|$).

Proof of Lemma 11

Let $\varepsilon > 0$ be given. Then $\frac{\varepsilon}{K}$ is also strictly greater than 0; so that by our hypothesis there exists an N such that

$$|s_n - l| < K \times \frac{\varepsilon}{K}$$
$$= \varepsilon \qquad \text{for all } n \geqslant N.$$

This is exactly the condition for $\{s_n\}$ to converge to l. ■

Remark

You will see just how useful this result is throughout the course!

Proof of Theorem 10

(i) Let $\varepsilon > 0$ be given.
Suppose first that $k \neq 0$. Then since $\lim_{n \to \infty} s_n = l_1$, there exists some N such that

$$|s_n - l_1| < \varepsilon \qquad \text{for all } n \geqslant N;$$

and so, multiplying through by $|k|$,

$$|ks_n - kl_1| < |k|\varepsilon \qquad \text{for all } n \geqslant N.$$

It follows from the "$K\varepsilon$ Lemma" with $K = |k|$ that $\lim\limits_{n\to\infty} (ks_n) = kl_1$.

On the other hand, if $k = 0$, then, for every ε and N,

$$|ks_n - kl_1| = 0 < \varepsilon \qquad \text{for all } n \geqslant N;$$

so that in this case also, $\lim\limits_{n\to\infty} (ks_n) = kl_1$.

(ii) Let $\varepsilon > 0$ be given.

Then we know from the definition of convergence of a sequence that there exist natural numbers N_1, N_2 such that

$$\left.\begin{aligned} &|s_n - l_1| < \varepsilon \qquad \text{for all } n \geqslant N_1, \text{ and} \\ &|t_n - l_2| < \varepsilon \qquad \text{for all } n \geqslant N_2. \end{aligned}\right\} \quad (1)$$

Since we are interested in s_n and t_n only for large indices n, we may as well suppose that n is greater than or equal to both N_1 and N_2. In other words,

$$|s_n - l_1| < \varepsilon \quad \text{and} \quad |t_n - l_2| < \varepsilon \qquad \text{for all } n \geqslant \max(N_1, N_2).$$

To show that $\{s_n + t_n\}$ converges to $l_1 + l_2$, we have to look at the difference $|(s_n + t_n) - (l_1 + l_2)|$. So, for all $n \geqslant \max(N_1, N_2)$,

$$\begin{aligned} |(s_n + t_n) - (l_1 + l_2)| &= |(s_n - l_1) + (t_n - l_2)| && \text{(by rearrangement)} \\ &\leqslant |s_n - l_1| + |t_n - l_2| && \text{(by the Triangle Inequality)} \\ &< \varepsilon + \varepsilon && (\text{as } n \geqslant N_1 \text{ and } n \geqslant N_2) \\ &= 2\varepsilon. \quad (2) \end{aligned}$$

The $K\varepsilon$ Lemma with $K = 2$ then shows that $\{s_n + t_n\}$ converges to $l_1 + l_2$.

To prove this result without using the $K\varepsilon$ Lemma, we must arrive at the following (rather than (2)):

$$|(s_n + t_n) - (l_1 + l_2)| < \varepsilon \qquad \text{for all } n \geqslant N \qquad (3)$$

for some N. To do this we should have chosen different N_1 and N_2 in (1) so that $|s_n - l_1| < \frac{1}{2}\varepsilon$ and $|t_n - l_2| < \frac{1}{2}\varepsilon$ for n larger than (the new values of) N_1 and N_2. If you follow through the rest of the proof you will see that you do end up with (3) rather than (2).

Of course, there is no way of knowing in advance that we should work with $\frac{1}{2}\varepsilon$ in (1)—so that the Lemma saves us going through the argument twice.

(iii) Let $\varepsilon > 0$ be given.

Then, as in (ii), we know that N_1 and N_2 exist such that (1) holds. To look at the convergence of $\{s_n t_n\}$ to $l_1 l_2$, we clearly have to look at the difference $|(s_n t_n) - (l_1 l_2)|$ for large n. Thus, for all $n \geqslant \max(N_1, N_2)$ (chosen so that both inequalities in (1) above hold),

$$\begin{aligned} |s_n t_n - l_1 l_2| &= |s_n(t_n - l_2) + l_2(s_n - l_1)| && \text{(a cunning rearrangement!)} \\ &\leqslant |s_n|\,|t_n - l_2| + |l_2|\,|s_n - l_1| && \text{(by the Triangle Inequality)} \\ &\leqslant |s_n|\varepsilon + |l_2|\varepsilon. \end{aligned}$$

This upper bound involves n, and so is not of the required form: (constant) × ε. Therefore we must use the result of Theorem 9 that, since $\{s_n\}$ is convergent, there exists some constant $K > 0$ such that $|s_n| < K$ for all n. It follows that, for all $n \geqslant \max(N_1, N_2)$, we have

$$\begin{aligned} |s_n t_n - l_1 l_2| &< K\varepsilon + |l_2|\varepsilon \\ &= (K + |l_2|)\varepsilon. \end{aligned}$$

The $K\varepsilon$ Lemma, with K replaced by $K + |l_2|$, shows that $s_n t_n \to l_1 l_2$ as $n \to \infty$. ■

Self-assessment Question

Prove that if the sequence $\{s_n\}$ converges to l, where each $s_n \neq 0$ and $l \neq 0$, then the sequence $\left\{\dfrac{1}{s_n}\right\}$ converges to $\dfrac{1}{l}$. (Hint: You may assume, without proof, that a constant K exists such that $|s_n| \geq K > 0$ for all n.)

Solution

Let $\varepsilon > 0$ be given.

We know that we have to look at the difference $\left|\dfrac{1}{s_n} - \dfrac{1}{l}\right|$; if we rewrite this in the form $\dfrac{|s_n - l|}{|s_n|\,|l|}$ and remember that we are looking for an upper bound for this expression, this gives us a guide as to how to proceed.

Since $s_n \to l$, there exists a natural number N such that $|s_n - l| < \varepsilon$ for all $n \geqslant N$. Hence, for all $n \geqslant N$,

$$\begin{aligned}\left|\frac{1}{s_n} - \frac{1}{l}\right| &= \frac{|s_n - l|}{|s_n|\,|l|}\\ &< \frac{\varepsilon}{|s_n|\,|l|} \quad \text{(as described above)}\\ &< \left(\frac{1}{K|l|}\right)\varepsilon \quad \text{(using the assumption).}\end{aligned}$$

The $K\varepsilon$ Lemma, with K replaced by $\dfrac{1}{K|l|}$, shows that $\dfrac{1}{s_n} \to \dfrac{1}{l}$ as $n \to \infty$.

2.3 Nested Intervals Theorem

Probably the main theorem in *M332* is the Cauchy–Goursat Theorem on page 63 of *Unit 5*. One of the key tools in its proof is a two-dimensional version of our next result. It is amazing how many of the previous results in the unit are required in order to prove it.

Theorem 12 (Nested Intervals Theorem)

Let $\{S_n\}$ be a sequence of nested closed intervals (that is, $S_n = [a_n, b_n]$ for some a_n, b_n, and $S_{n+1} \subseteq S_n$), and let their lengths, $l(S_n) = b_n - a_n$, converge to 0. Then

(i) $S = \bigcap_{n=1}^{\infty} S_n$ is non-empty and contains a unique real number x_0; and

(ii) given $\varepsilon > 0$, there is a natural number N such that

$$S_n \subseteq \{x : |x - x_0| < \varepsilon\} \qquad \text{for all } n \geqslant N.$$

(The second part of Theorem 12 tells us that we can always ensure that "eventually" all the intervals (or, if you prefer "all but a finite number") lie inside an arbitrarily small open interval at x_0.)

Proof of Theorem 12

(i) Since the S_n are nested, the sequence $\{a_n\}$ is necessarily increasing and also necessarily bounded above (by b_1, for example). It follows from Theorem 8 that the sequence $\{a_n\}$ converges to some real number $l_1 = \sup\{a_1, a_2, \ldots\}$; we shall show that $l_1 \in S$ and then that $\{l_1\} = S$.

Fix any $m \in \mathbf{N}$. Then, since the sequence $\{S_n\}$ is nested, it follows that $a_n \leqslant b_m$ for every n. Hence b_m is an upper bound for the sequence $\{a_n\}$; and so by the "leastness" of l_1 as an upper bound for the sequence, $l_1 \leqslant b_m$.

Since $\{a_n\}$ is increasing, $a_n \leqslant l_1$ for each n; and, as we have just seen, $l_1 \leqslant b_n$ for each n. Consequently $l_1 \in [a_n, b_n] = S_n$, for each n; in other words,

$$l_1 \in S = \bigcap_{n=1}^{\infty} S_n.$$

In the same way, we can check that the sequence b_n decreases and converges to its infimum l_2, and that $l_2 \in S$. To complete the proof of (i) we need only show that $l_1 = l_2$.

For $[l_1, l_2] \subseteq S$, and no larger interval can lie in S, by the definitions of l_1 and l_2.

As $n \to \infty$, we have $l(S_n) \to 0$, $a_n \to l_1$ and $b_n \to l_2$. Now let $n \to \infty$ in the identity

$$b_n - a_n = l(S_n)$$

and use parts (i) and (ii) of Theorem 10 (with $k = -1$). It follows that

$$l_2 - l_1 = 0,$$

or

$$l_1 = l_2.$$

Thus the x_0 in (i) is the common value of l_1 and l_2.

(ii) Let $\varepsilon > 0$ be given.

Since $x_0 - \varepsilon$ is less than the limit of the increasing sequence $\{a_n\}$, there exists a natural number N_1 such that

$$x_0 - \varepsilon < a_n \leqslant x_0 \qquad \text{for all } n \geqslant N_1;$$

and since $x_0 + \varepsilon$ is greater than the limit of the decreasing sequence $\{b_n\}$, there exists a natural number N_2 such that

$$x_0 \leqslant b_n < x_0 + \varepsilon \qquad \text{for all } n \geqslant N_2.$$

Hence if we put $N = \max(N_1, N_2)$, it follows that

$$x_0 - \varepsilon < a_n \leqslant x_0 \leqslant b_n < x_0 + \varepsilon \qquad \text{for all } n \geqslant N;$$

in other words,

$$[a_n, b_n] = S_n \subseteq (x_0 - \varepsilon, x_0 + \varepsilon) \qquad \text{for all } n \geqslant N. \quad ■$$

Self-assessment Question

For each of the following sequences of intervals $\{S_n\}$, state whether the conclusions (i) and (ii) of Theorem 12 hold. Give reasons for your assertions, but you need not prove your assertions in detail.

(a) $S_n = \left[0, \frac{1}{n}\right]$

(b) $S_n = \left(0, \frac{1}{n}\right)$

(c) $S_n = [(-1)^n, (-1)^n + 2]$

(d) $S_n = ((-1)^n, (-1)^n + 2)$

Solution

(a) Both (i) and (ii) hold (with $x_0 = 0$), as the hypotheses of Theorem 12 are satisfied.

(b) Here $\bigcap_{n=1}^{\infty} S_n = \varnothing$, so that (i) does not hold; however, the conclusion (ii) holds with $x_0 = 0$.

(c) Since

$$S_n = \begin{cases} [-1, 1], & \text{when } n \text{ is odd,} \\ [1, 3], & \text{when } n \text{ is even,} \end{cases}$$

$\bigcap_{n=1}^{\infty} S_n$ is the single point 1, so that (i) holds (with $x_0 = 1$).

However, (ii) does not hold. For example, if we take $\varepsilon = 1$, it is not true that eventually all S_n lie in

$$(x_0 - \varepsilon, x_0 + \varepsilon) = (0, 2).$$

(d) Since

$$S_n = \begin{cases} (-1, 1), & \text{when } n \text{ is odd,} \\ (1, 3), & \text{when } n \text{ is even,} \end{cases}$$

$\bigcap_{n=1}^{\infty} S_n = \varnothing$, so that (i) fails to hold.

Again (ii) cannot hold. For example, for any choice of x_0, if we take $\varepsilon = \frac{1}{2}$, it cannot be true that eventually all S_n lie in the interval $(x_0 - \frac{1}{2}, x_0 + \frac{1}{2})$; each S_n is "too long".

2.4 Sequences of Functions

Let $f_n(x)$, $n = 1, 2, 3, \ldots$, be a sequence of functions defined on some set $A \subseteq \mathbf{R}$ (usually A will be an interval). We are going to look at the possible convergence of the functions $\{f_n\}$ to some limit function f in two different ways, *pointwise* and *uniformly*.

Definition

The sequence of functions $\{f_n\}$ **converges pointwise** to (the limit function) f on a set $A \subseteq \mathbf{R}$ if

$$\lim_{n\to\infty} f_n(x) = f(x) \qquad \text{for each } x \in A.$$

In other words, given any $\varepsilon > 0$ and any $x \in A$, there exists a corresponding natural number N such that

$$|f_n(x) - f(x)| < \varepsilon \qquad \text{for all } n \geqslant N;$$

here N may depend on the sequence of functions, on ε and on x.

Example

Show that the sequence of functions $\{f_n\}$ given by

$$f_n(x) = \frac{nx}{x+n}$$

converges pointwise on $(-1, 1)$.

Solution

The first step is to find a suitable function f that can then be shown to be the limit function of the sequence $\{f_n(x)\}$.

We are interested in the value of $f_n(x)$ only for large indices n; since we can write $f_n(x)$ in the form $\dfrac{x}{\frac{x}{n}+1}$, and the denominator approaches the value 1 as n gets larger and larger, it looks as though the function $f(x) = x$ is a good candidate as a limit function. We now check that this is the case, from the above definition.

We know that we should look at the difference $|f_n(x) - f(x)| = |f_n(x) - x|$. So let $\varepsilon > 0$ be given. Then

$$\begin{aligned} |f_n(x) - x| &= \left| \frac{nx}{x+n} - x \right| \\ &= \frac{x^2}{|x+n|} \\ &< \frac{1}{|x+n|} && \text{(since } |x| < 1\text{)} \\ &\leqslant \frac{1}{n - |x|} && \text{(since } |x+n| \geqslant ||x| - |n|| = n - |x|\text{, by the} \\ & && \text{"cunning form" of the Triangle Inequality)} \\ &< \varepsilon && \text{if } n \geqslant N, \end{aligned}$$

where we make the choice that N is the least natural number with $N > \frac{1}{\varepsilon} + |x|$. Hence the sequence $\{f_n\}$ converges to the identity function pointwise on $(-1, 1)$.

Example

Show that $\lim\limits_{n \to \infty} \frac{x^n}{n} = 0$ pointwise on $(-1,1)$.

Solution

Put $f_n(x) = \frac{x^n}{n}$ and let $g_n(x) = |f_n(x)|$, where $x \in (-1, 1)$.

Since $g_{n+1}(x) = \frac{|x|^{n+1}}{n+1}$ and $g_n(x) = \frac{|x|^n}{n}$, we see that

$$\begin{aligned} g_{n+1}(x) &= |x| \frac{1}{1 + (1/n)} g_n(x) && (1) \\ &< g_n(x) && \text{(as } |x| < 1 \text{ and } n \geqslant 1\text{).} \end{aligned}$$

Thus the sequence $\{g_n(x)\}$ is decreasing. It is also bounded below, by 0; hence, by Theorem 8, $\{g_n(x)\}$ is convergent to some limit l. Letting $n \to \infty$ on both sides of (1) we deduce that $l = |x|l$, so that $l(1 - |x|) = 0$. Since $|x| \neq 1$, it follows that necessarily $l = 0$.

Now let $\varepsilon > 0$ be given. Since, for any given x, the sequence $\{g_n(x)\}$ converges to 0 as $n \to \infty$, it follows that there exists some natural number N such that

$$|g_n(x) - 0| < \varepsilon \qquad \text{for all } n \geqslant N. \qquad (2)$$

Hence we deduce immediately that, for all $n \geqslant N$,

$$\begin{aligned} |f_n(x) - 0| &= |g_n(x) - 0| && \text{(from the definitions of } f_n \text{ and } g_n\text{)} \\ &< \varepsilon && \text{(by (2));} \end{aligned}$$

in other words, $\lim\limits_{n \to \infty} f_n(x) = 0$ as desired.

Remark

For further reference, we mention without proof that

$$\lim_{n \to \infty} \frac{x^n}{n!} = 0 \qquad \text{pointwise on } \mathbf{R}.$$

Self-assessment Question

Show that the sequence of functions $\{f_n\}$ given by $f_n(x) = \dfrac{\sin(n^2 x)}{n^2 + x}$ converges pointwise on $(-1, 1)$.

Solution

Since the numerator in the expression for $f_n(x)$ is bounded (by 1) and the denominator gets larger as n gets larger, it looks reasonably likely that the sequence converges pointwise to the zero function on $(-1, 1)$.

So, let $\varepsilon > 0$ be given. Then

$$\begin{aligned} |f_n(x) - 0| &= \left|\frac{\sin(n^2 x)}{n^2 + x}\right| \\ &\leqslant \frac{1}{|n^2 + x|} \\ &\leqslant \frac{1}{n^2 - |x|} && \text{(by the Triangle Inequality)} \\ &< \varepsilon && \text{for all } n \geqslant N, \end{aligned}$$

if we make the choice that N is the least natural number with $N > \left(\dfrac{1}{\varepsilon} + |x|\right)^{\frac{1}{2}}$.

Hence $\{f_n\}$ converges pointwise (to the zero function) on $(-1, 1)$.

It is very tempting to believe that, if a sequence $\{f_n\}$ converges pointwise to f, then the limit function f should inherit whatever "nice" properties each of the f_n might have. For example, if each f_n were continuous then maybe f is also continuous. However this is *not* necessarily the case.

However, if we require the functions f_n to satisfy a stronger condition than pointwise convergence, it turns out that f will inherit many of the "nice" properties of the f_n.

Definition

The sequence of functions $\{f_n\}$ **converges uniformly** to (the limit function) f on a set $A \subseteq \mathbf{R}$ if given any $\varepsilon > 0$ there exists a corresponding N such that

$$|f_n(x) - f(x)| < \varepsilon \qquad \text{for all } n \geqslant N \text{ and all } x \in A;$$

here N may depend on the sequences and on ε, but the *same* N must serve for *all values of* x in A.

Notice that in the Example on page 30, corresponding to any $\varepsilon > 0$, if we choose N to be the least natural number greater than $\dfrac{1}{\varepsilon} + 1$, then

$$|f_n(x) - x| < \varepsilon \qquad \text{for all } n \geqslant N \text{ and all } x \in (-1, 1).$$

Hence $\{f_n(x)\}$ is also uniformly convergent to x on $(-1, 1)$.

We shall return to uniform convergence, in **C**, in detail in *Unit 11*, where we shall generalize the following two results for real functions (which we state without proof).

Properties of uniform convergence

If the sequence $\{f_n\}$ is uniformly convergent to f on an interval $[a, b]$,

(i) if each f_n is continuous, so is f;

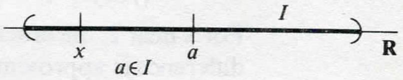
(ii) if $\int_a^b f_n$ and $\int_a^b f$ exist, then $\lim\limits_{n\to\infty} \int_a^b f_n = \int_a^b f$.

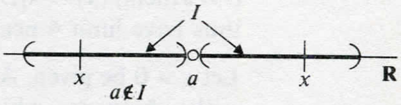
We use the abbreviated notation for $\int_a^b f_n(x)dx$, etc.

3 LIMITS OF FUNCTIONS, CONTINUITY

In this section we shall make precise the intuitive concept that "if x is near some point a then the corresponding value of $f(x)$ must be near some value l". We shall look at the statement in two closely related ways, one where we do not care whether f is even defined at a and one where the value of $f(a)$ plays a key role.

3.1 Limits





Suppose that the function f has a domain which contains an open interval I around a (finite) point a, except that we do not mind whether or not a itself belongs to the domain. If, for all choices of a point x (except $x = a$) in I, $f(x)$ gets closer and closer to a finite number l as x gets closer and closer to a, we call l the limit of f near a. It is common to call l the limit "at a" rather than "near a", but this is rather misleading since in our definition of limit we are specifically ignoring what (if anything) is happening at a. We can rephrase this definition very precisely in terms of ε and δ:

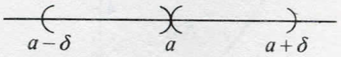
Definition

Let f be a function, and $a, l \in \mathbf{R}$. We say that f has **limit l near a** if, to every given $\varepsilon > 0$, there corresponds some $\delta > 0$ such that

$$\text{if } 0 < |x - a| < \delta \quad \text{then } |f(x) - l| < \varepsilon.$$

Sometimes we write this as "$\lim_{x \to a} f(x) = l$" or as "$f(x) \to l$ as $x \to a$", or say that "$f(x)$ tends to l as x tends to a".

Remark



Without making a fuss about it, we shall always suppose that δ is chosen sufficiently small that $\{x : 0 < |x - a| < \delta\}$ lies in I. Notice the vital inequality "$0 < |x - a|$" in the definition—this ensures that we do *not* take into account the point a.

In *M203 Analysis Block B, Unit 5, Proofs*, page 19 it is shown that $\lim_{x \to a} f(x) = l$ in the *M203* sense if and only if the above ε-δ definition is satisfied.

Just as in Section 2.2 (for sequences), we notice that the ε which appears in the definition of limit can be replaced by "(a constant) × ε" without affecting the limiting behaviour of f.

Lemma 13 (Another "$K\varepsilon$ Lemma")

Suppose that, corresponding to each given $\varepsilon > 0$, there exists a $\delta > 0$ such that

$$|f(x) - l| < K\varepsilon \qquad \text{whenever } 0 < |x - a| < \delta,$$

where K is a (positive) real number independent of ε and δ. Then f has limit l near a.

Example

(a) Show that the function

$$f : x \longrightarrow \begin{cases} x^2, & \text{if } x \neq 2, \\ 0, & \text{if } x = 2, \end{cases}$$

has limit 4 as x tends to 2.

(b) Show that the function

$$g: x \longrightarrow \begin{cases} 1 + x, & \text{if } x > 0, \\ -1 + x, & \text{if } x < 0, \end{cases}$$

does not tend to a limit as x tends to 0.

Solution

(a) From the definition, we know that we are supposed to show that for x near 2 the following difference is small:

$$|f(x) - 4| = |x^2 - 4| = |(x - 2)(x + 2)|.$$

For x near 2, the bracket $(x + 2)$ is more or less just the number 4, and so the difference is approximately $4|x - 2|$. So if $0 < |x - 2| < \delta$ for some choice of δ, then $|f(x) - 4|$ is less than about 4δ; for example, if we make the choice $\delta = \varepsilon$ then $|f(x) - 4|$ is less than about 4ε, so that by Lemma 13 the function will thus have limit 4 near 2. We now make this outline argument precise.

Let $\varepsilon > 0$ be given. As we are interested only in x near 2, consider to begin with only those x which lie in the open interval $(1, 3)$. For these, $x + 2 \in (3, 5)$, so that $|x + 2| < 5$. Hence if we make the choice

$$\delta = \min(1, \varepsilon)$$

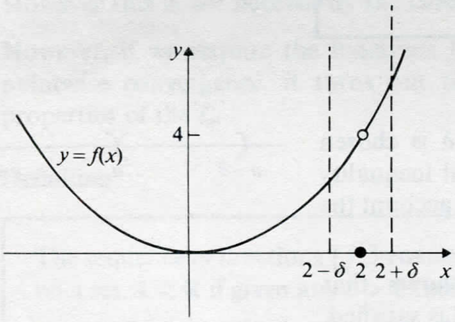
(chosen so that necessarily x is near to 2 and $|x - 2|$ is small if ε is small), then for all x for which $0 < |x - 2| < \delta$ we have

$$|x - 2| < \varepsilon \quad \textit{and} \quad |x + 2| < 5,$$

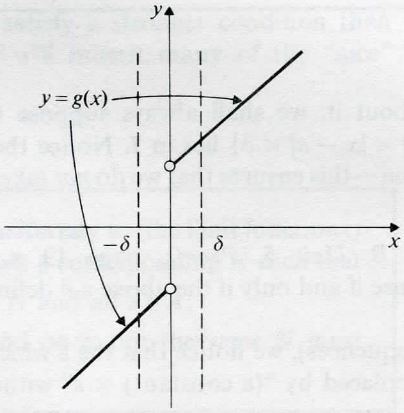
so that

$$|f(x) - 4| = |x - 2|\,|x + 2| < 5\varepsilon.$$

By the $K\varepsilon$ Lemma (Lemma 13), it follows that $f(x) \to 4$ as $x \to 2$.



(a)



(b)

(b) Suppose that some real number l exists such that $\lim\limits_{x \to 0} g(x) = l$. Then, corresponding to any given $\varepsilon > 0$, there exists a $\delta > 0$ such that

$$|g(x) - l| < \varepsilon \qquad \text{whenever } 0 < |x - 0| < \delta.$$

For such x, $g(x)$ must belong to the interval $(l - \varepsilon, l + \varepsilon)$ of length 2ε; if ε is less than $\frac{1}{2}$, so that the interval has length less than 1, it cannot contain values near both 1 and -1, although near 0, g is taking values near both. This suggests that if we choose an ε less than $\frac{1}{2}$ we should be able to get a contradiction. We now write out the formal argument.

Take $\varepsilon = \dfrac{1}{4}$. Then if the limit $l = \lim\limits_{x \to 0} g(x)$ exists there must exist a corresponding $\delta > 0$ such that

$$|g(x) - l| < \frac{1}{4} \qquad \text{whenever } 0 < |x| < \delta.$$

Taking $x = \tfrac{1}{2}\delta$, $x = -\tfrac{1}{2}\delta$ in turn, so that $g(\tfrac{1}{2}\delta) = 1 + \tfrac{1}{2}\delta$ and $g(-\tfrac{1}{2}\delta) = -1 - \tfrac{1}{2}\delta$, we see that the following two inequalities must *both* be satisfied:

$$|1 + \tfrac{1}{2}\delta - l| < \tfrac{1}{4} \quad \text{and} \quad |1 + \tfrac{1}{2}\delta + l| < \tfrac{1}{4}.$$

For $|-1 - \tfrac{1}{2}\delta - l| = |1 + \tfrac{1}{2}\delta + l|$.

Using the Triangle Inequality (as we did in a similar situation in the Example on page 24),

$$|(1 + \tfrac{1}{2}\delta - l) + (1 + \tfrac{1}{2}\delta + l)| \leqslant |1 + \tfrac{1}{2}\delta - l| + |1 + \tfrac{1}{2}\delta + l|$$

so that

$$\begin{aligned} 2 + \delta &< \frac{1}{4} + \frac{1}{4} \\ &= \frac{1}{2}. \end{aligned}$$

This is a contradiction; consequently $\lim_{x \to 0} g(x)$ does not exist.

Self-assessment Question

Let f be the function

$$f(x) = \begin{cases} x^{\frac{1}{2}}, & \text{if } x > 0, \\ 0, & \text{if } x < 0, \\ 1, & \text{if } x = 0. \end{cases}$$

Prove that:

(a) $\lim_{x \to a} f(x) = a^{\frac{1}{2}}$ for each $a > 0$;

(b) $\lim_{x \to 0} f(x) = 0$.

(Hint: $(x^{\frac{1}{2}} - a^{\frac{1}{2}})(x^{\frac{1}{2}} + a^{\frac{1}{2}}) = x - a$ when $x, a \geqslant 0$.)

Solution

(a) We know that we have to examine the difference $|f(x) - a^{\frac{1}{2}}|$. If $a > 0$ and x is near a (so that $x > 0$) then

$$\begin{aligned} |f(x) - a^{\frac{1}{2}}| &= |x^{\frac{1}{2}} - a^{\frac{1}{2}}| \\ &= \frac{|x - a|}{x^{\frac{1}{2}} + a^{\frac{1}{2}}} \quad \text{(denominator non-zero).} \end{aligned}$$

As x is near to a, this ratio is approximately $\dfrac{1}{2a^{\frac{1}{2}}}|x - a|$, and is certainly less than or equal to $\dfrac{1}{a^{\frac{1}{2}}}|x - a|$. It is now clear how to proceed.

Let $\varepsilon > 0$ be given. Taking $\delta = \min(\varepsilon, \tfrac{1}{2}a)$ (chosen to make sure that $x > 0$), we see that if $0 < |x - a| < \delta$ $(\leqslant \varepsilon)$, then

$$\begin{aligned} |f(x) - a^{\frac{1}{2}}| &= \frac{|x - a|}{x^{\frac{1}{2}} + a^{\frac{1}{2}}} \\ &< \frac{|x - a|}{a^{\frac{1}{2}}} < a^{-\frac{1}{2}}\varepsilon. \end{aligned}$$

The $K\varepsilon$ Lemma (with $K = a^{-\frac{1}{2}}$) shows that $f(x) \to a^{\frac{1}{2}}$ as $x \to a$.

(b) Here, for x near to 0 but not equal to 0,

$$|f(x) - 0| = \begin{cases} x^{\frac{1}{2}}, & \text{if } x > 0, \\ 0, & \text{if } x < 0. \end{cases}$$

In either case,

$$|f(x) - 0| < |x|^{\frac{1}{2}}.$$

This will be less than ε if we arrange that $|x| < \varepsilon^2$; so choosing $\delta = \varepsilon^2$ may well work. Let us try it.

As usual, let $\varepsilon > 0$ be given. Put $\delta = \varepsilon^2$. Then, for all x with $0 < |x - 0| < \delta$, it follows that

$$\begin{aligned} |f(x) - 0| &< |x|^{\frac{1}{2}} \\ &< \delta^{\frac{1}{2}} = \varepsilon. \end{aligned}$$

This proves that $\lim\limits_{x \to 0} f(x)$ exists and equals 0.

Self-assessment Question

Let $f: \mathbf{R} \longrightarrow \mathbf{R}$ be a function, and $a \in \mathbf{R}$. Which of the following statements, if any, are equivalent to the statement that $\lim\limits_{x \to a} f(x)$ exists and is equal to l? Justify your answers.

(a) For each given $\delta > 0$, there exists $\varepsilon > 0$ such that

$$|f(x) - l| < \varepsilon \qquad \text{for every } x \text{ such that } 0 < |x - a| < \delta.$$

(b) For each given $\delta > 0$, there exists $\varepsilon > 0$ such that

$$|f(x) - l| < \delta \qquad \text{for every } x \text{ such that } 0 < |x - a| < \varepsilon.$$

(c) For each given $\varepsilon > 0$, there exists $\delta > 0$ such that

$$|f(x) - l| < \varepsilon \quad \text{implies that} \quad 0 < |x - a| < \delta.$$

Solution

(a) This is not equivalent to the definition of "$\lim\limits_{x \to a} f(x) = l$".

The function

$$f(x) = \begin{cases} 1 + x, & \text{if } x > 0, \\ -1 + x, & \text{if } x < 0 \end{cases}$$

(which we met in Example (b) on page 34) will do as a counter-example. For condition (a) is satisfied for $a = 0$ with $l = 0$, by taking $\varepsilon = \delta + 1$ for example. However, as we saw on page 34, $\lim\limits_{x \to 0} f(x)$ does not exist.

(b) This is precisely the definition of $\lim\limits_{x \to a} f(x) = l$. The interchange of the symbols ε and δ is of no significance—it is just a mean trick on our part!

(c) This is not equivalent to the definition of limit.

For example, let $a = l = 0$ and let f: $\mathbf{R} \longrightarrow \mathbf{R}$ be the constant function $f: x \longrightarrow 0$. Then certainly $\lim\limits_{x \to 0} f(x) = 0$ (as you can easily check); and also,

$$|f(x) - l| = 0 \qquad \text{for every } x, \text{ however large } x \text{ may be.}$$

Thus, given $\varepsilon > 0$ it is always true that $|f(x) - l| < \varepsilon$; and there is no restriction on x, hence no $\delta > 0$ such that condition (c) is satisfied.

The idea of limit is particularly useful in our discussions of continuity and differentiability. This interplay is a two-way affair; for example l'Hôpital's Rule (see page 52) uses differentiation to evaluate certain types of limits; and the fact that $\dfrac{\sin x}{x} \longrightarrow 1$ as $x \longrightarrow 0$ is a crucial step in finding the derivative of $\sin x$.

Just as we were able to prove (on page 26) some results for combinations of convergent sequences, similar results hold for limits of functions; we state these (without proof) as

Theorem 14

Let $f(x) \longrightarrow l_1$, $g(x) \longrightarrow l_2$ as $x \longrightarrow a$, and let $k \in \mathbf{R}$. Then, as $x \longrightarrow a$,

(a) $kf(x) \longrightarrow kl_1$,
(b) $f(x) + g(x) \longrightarrow l_1 + l_2$, and
(c) $f(x)g(x) \longrightarrow l_1 l_2$.

(The proofs are very similar to those of Theorem 10 on page 26.)

Self-assessment Question

Let $f(x) \to l_1$ as $x \to a$ and $g(x) \to l_2$ as $x \to l_1$. By means of a specific counter-example, show that it is *not* necessarily true that

$$g \circ f(x) \to l_2 \text{ as } x \to a.$$

(Hint: Read the definition of limit carefully!)

Solution

Let $f(x) = 0$ for all x, $l_1 = a = 0$.

Next, let

$$g(x) = \begin{cases} 1, & \text{if } x \neq 0, \\ 2, & \text{if } x = 0. \end{cases}$$

From the definition of limit, we see that $g(x) \to 1$ as $x \to 0$, so that $l_2 = 1$; and since $g \circ f(x) = 2$ for each $x \in \mathbf{R}$, $\lim\limits_{x \to 0} (g \circ f)(x) = 2$ although $l_2 = 1$.

Remark

The key point here is the condition $0 < |x - a|$ in the definition of limit on page 33.

3.2 More Limits

In the previous section we looked at the definition of $\lim\limits_{x \to a} f(x) = l$ for finite values of a and of l, under the hypothesis that the domain of f contained some open interval around a (except that a itself might not belong to the interval).

As x approaches $+\infty$, we would mean by the statement "$\lim\limits_{x \to +\infty} f(x) = l$" that as x gets larger and larger, $f(x)$ gets closer and closer to the value l. In terms of ε, we can rephrase this as

Definition

Let f be a function, and $l \in \mathbf{R}$. We say that f has **limit** l **near** $+\infty$ if, to every given $\varepsilon > 0$, there corresponds some real number M such that

if $x > M$ then $|f(x) - l| < \varepsilon$.

(Of course we always think of M as a "large positive number".)

In the same way, as x approaches a finite number a, we would mean by the statement "$\lim\limits_{x \to a} f(x) = +\infty$" that as x gets closer and closer to a, the value of $f(x)$ gets larger and larger. We can express this idea in a precise way as

Definition

Let f be a function. We say that f has **limit** $+\infty$ **near** a if, to every given M, there corresponds some $\delta > 0$ such that

if $0 < |x - a| < \delta$ then $f(x) > M$.

Again we exclude the point $x = a$ from consideration.

You should now turn back to page 33 to compare these definitions with the definition of $\lim\limits_{x \to a} f(x) = l$ where a, l are both finite; notice how we have modified that definition to allow a, l to become $+\infty$ in turn.

Example

Show that the function $f(x) = \dfrac{x}{x+1}$ $(x \neq -1)$ tends to a limit as $x \to +\infty$.

Solution

Let us rewrite the formula for $f(x)$ in the form

$$f(x) = \frac{1}{1 + \dfrac{1}{x}}.$$

Then as x gets larger and larger, $\dfrac{1}{x}$ gets closer and closer to 0, so that $\dfrac{1}{1+\dfrac{1}{x}}$ approaches the value 1. Hence it seems reasonable to guess that $f(x) \to 1$ as $x \to \infty$. We now prove this carefully.

We shall deal only with $x > -1$, as we are interested in the behaviour of $f(x)$ only for large x.

We know that we should look at the expression

$$\begin{aligned}|f(x) - 1| &= \left|\frac{x}{x+1} - 1\right| \\ &= \frac{1}{|x+1|} \\ &= \frac{1}{x+1} \qquad \text{(so long as } x > -1\text{)},\end{aligned}$$

and this will be less than ε if we arrange that $x + 1 > \dfrac{1}{\varepsilon}$.

So we have two conditions to satisfy, namely $x > -1$ and $x > -1 + \dfrac{1}{\varepsilon}$; in fact the condition that $x > -1 + \dfrac{1}{\varepsilon}$ includes the other. This argument suggests what choice we might make of M in the definition.

So, let $\varepsilon > 0$ be given. Then choose $M = -1 + \dfrac{1}{\varepsilon}$, so that for all $x > M$,

$$|f(x) - 1| = \frac{1}{|x+1|} < \varepsilon.$$

It follows that $f(x) \to 1$ as $x \to +\infty$.

Remarks

We now make some observations about the two definitions that are of real use in practice.

(i) It is enough in the first definition to prove that corresponding to any given ε we can find an M such that

$$\text{if } x > M \quad \text{then } |f(x) - l| < K\varepsilon$$

for some fixed constant K.

(ii) In the second definition it is quite enough to suppose that we can find a δ corresponding to any given value of M that is larger than some fixed value K. For if we can find a $\delta > 0$ such that

$$0 < |x - a| < \delta \quad \text{implies that} \quad f(x) > M$$

for those values of M greater than K, the same δ will do for all smaller values of M.

Example

Show that the function $f(x) = 1/(x+1)^2$ tends to $+\infty$ as $x \to -1$.

Solution

As we explained above, we may, without loss of generality, consider only positive values of M in our discussion, i.e. we assume throughout that $M > 0$. (This is the case $K = 0$ in part (ii) of the above Remarks.)

We have that $f(x) > M$, that is $1/(x+1)^2 > M$, whenever $(x+1)^2 < \frac{1}{M}$; and this is the same condition as $|x+1| < \frac{1}{M^{\frac{1}{2}}}$. So it looks as though the choice $\delta = M^{-\frac{1}{2}}$ would be a good one.

Let $M > 0$ be given. Choose $\delta = M^{-\frac{1}{2}}$. Then we have, for all x such that $0 < |x - (-1)| < \delta$,

$$\begin{aligned} f(x) &= \frac{1}{(x+1)^2} \\ &= \frac{1}{(x-(-1))^2} \\ &> \frac{1}{\delta^2} = M. \end{aligned}$$

It follows that $f(x) \to +\infty$ as $x \to -1$, as required.

Self-assessment Question

Show that, for each positive integer m, the function

$$f(x) = \frac{x^m}{e^x}$$

tends to 0 as x tends to $+\infty$.

(Hint: Use the fact that, for positive values of x,

$$e^x = 1 + x + \frac{x^2}{2!} + \ldots + \frac{x^{m+1}}{(m+1)!} + \ldots > \frac{x^{m+1}}{(m+1)!}\text{)}$$

Solution

We are told that m is a fixed integer and that $e^x > \frac{x^{m+1}}{(m+1)!}$ for all $x > 0$. We know from the definition that we should look at the difference

$$\begin{aligned} |f(x) - 0| &= \left|\frac{x^m}{e^x}\right| \\ &= \frac{x^m}{e^x} \qquad \text{(for positive } x\text{).} \end{aligned}$$

If we rewrite the inequality in the hint in the form

$$\frac{x^m}{e^x} < \frac{(m+1)!}{x},$$

we see that we can make $|f(x) - 0|$ less than ε by ensuring that $\frac{(m+1)!}{x} < \varepsilon$. This suggests a choice of M in the limit definition.

Let $\varepsilon > 0$ be given. Choose $M = (m+1)!/\varepsilon$. Then for all x with $x > M$ we have that

$$\begin{aligned} |f(x) - 0| &= \frac{x^m}{e^x} < \frac{(m+1)!}{x} \\ &< \frac{(m+1)!}{M} = \varepsilon; \end{aligned}$$

in other words, $\lim_{x \to +\infty} f(x) = 0$, as required.

Remark

The result of this SAQ is used on page 9 of *Unit 8*.

Self-assessment Question

Show that the function $f(x) = \dfrac{e^x}{(x-2)^2}$ $(x \neq 2)$ tends to $+\infty$ as $x \to 2$.

Solution

As x approaches 2, e^x approaches the value e^2; meanwhile, $(x-2)^2$ approaches 0 $\quad (x-2)^2 \geqslant 0.$ so that $\dfrac{1}{(x-2)^2}$ approaches $+\infty$. Hence the product $e^x \cdot \dfrac{1}{(x-2)^2}$ will approach $+\infty$. Let us now make this precise.

Choose any $M > 0$. If we arrange to consider only positive values of x, then we shall have $e^x > 1$. If we additionally arrange to consider only those x for which $|x-2| < M^{-\frac{1}{2}}$, then we also have that $\dfrac{1}{(x-2)^2} > M$.

So let us choose $\delta = \min(2, M^{-\frac{1}{2}})$. Then for all x such that $0 < |x-2| < \delta$, all the inequalities of the previous paragraph hold, so that

$$\begin{aligned} f(x) &= e^x \cdot \frac{1}{(x-2)^2} \\ &> 1 \cdot M = M. \end{aligned}$$

It follows that $f(x) \to +\infty$ as $x \to 2$.

Many useful functions of a real variable are defined only on subsets of **R**; for example, $\log x$ makes sense only when $x > 0$. Consequently, we shall often need to deal with the limiting behaviour of a function as x approaches some point either through larger values of x only or through smaller values of x only.

Definition

Let f be a function defined on an open interval (a, b). Then f has **right-hand limit** l_1 **near** a if, to every given $\varepsilon > 0$, there corresponds some $\delta > 0$ such that

$$\text{if } a < x < a + \delta \quad \text{then } |f(x) - l_1| < \varepsilon;$$

and f has **left-hand limit** l_2 **near** b if, to every given $\varepsilon > 0$, there corresponds some $\delta > 0$ such that

$$\text{if } b - \delta < x < b \quad \text{then } |f(x) - l_2| < \varepsilon.$$

Sometimes we write these limits as $\lim\limits_{x \to a^+} f(x) = l_1$ and $\lim\limits_{x \to b^-} f(x) = l_2$.

We remark without proof that from the result of the SAQ on page 39 we can (with some effort) deduce that

$$\lim_{x \to 0^+} x \log x = 0.$$

This particular limit will be invaluable in our evaluation of certain integrals in *Unit 11*; for example, in line −15 on page 107.

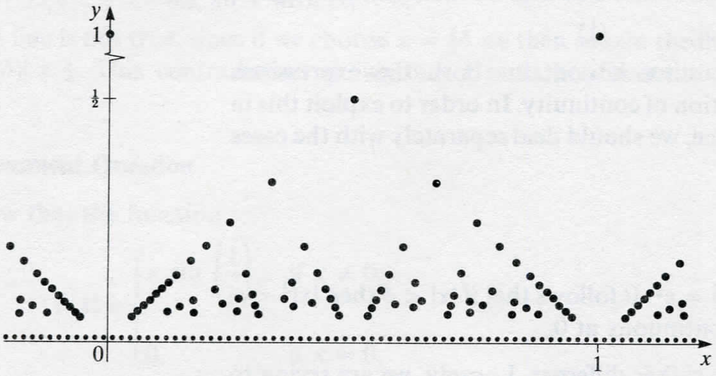
3.3 Continuity

An imprecise concept of continuity is that a function f is said to be *continuous* if its graph can be drawn as a continuous curve; that is, without lifting pen from paper. This is not a useful description of continuity, since there exist functions that are continuous at each irrational point but discontinuous at each rational point! One such function is *Riemann's function* on $\mathbf{R}$:

$$f(x) = \begin{cases} 1/q, & \text{if } x \text{ is a rational number } p/q, \quad \text{where} \quad q > 0, \\ 0, & \text{if } x \text{ is irrational.} \end{cases}$$

Here, we assume that p/q is in its lowest terms—that is, p and q have no common factor.

It has the following graph.



If the graph of f is to be, in some sense, unbroken at the point $(a, f(a))$, then we shall want all points $(x, f(x))$ to be near to $(a, f(a))$ when x is near to a. This will be the case whenever

x close to a *implies* $f(x)$ close to $f(a)$.

Recalling the definition of *limit* in Section 3.1, we see that we can rephrase this as

$\lim_{x \to a} f(x)$ exists and equals $f(a)$.

While this would do as a formal definition of continuity, for ease of application it is better to rewrite it in terms of the ε and δ that we used in the definition of limit.

Definition

The function f is **continuous at the point** a if, corresponding to every given $\varepsilon > 0$, there exists some $\delta > 0$ such that

$$\text{if } |x - a| < \delta \quad \text{then } |f(x) - f(a)| < \varepsilon.$$

We say that f is **continuous on an interval** I if it is continuous at each point of I.

Remark

Notice that the condition "$0 < |x - a|$" which was so important in the definition of a limit on page 33 does not appear in the definition of continuity here. For a limit at a we do not care whether or not the function f is even defined at the point a; however, it would not make sense to talk about continuity at a if we were to ignore the existence of some corresponding point $(a, f(a))$ on the graph of f.

As usual, before giving an example of our definition involving ε, we state (this time without proof) the corresponding "$K\varepsilon$ Lemma".

Lemma 15

Suppose that, corresponding to each given $\varepsilon > 0$, there is a $\delta > 0$ such that

$$|f(x) - f(a)| < K\varepsilon \qquad \text{for all } x \text{ satisfying } |x - a| < \delta,$$

where K is a (positive) real number independent of ε and x. Then f is continuous at a.

Example

Prove that the function $f: \mathbf{R} \longrightarrow \mathbf{R}$ defined by $f: x \longrightarrow x^2$ is continuous at each point $a \in \mathbf{R}$.

Solution

From the definition, we know that we should consider the difference between $f(x)$ and $f(a)$:

$$|f(x) - f(a)| = |x^2 - a^2|.$$

The first step is to notice that the right-hand side can be factorized, so that

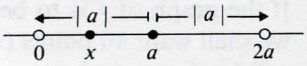
$$|f(x) - f(a)| = |x - a|\,|x + a|; \qquad (1)$$

this is obviously a useful formula since it contains both the expressions $|f(x) - f(a)|$ and $|x - a|$ in the definition of continuity. In order to exploit this in detail, it turns out that, for convenience, we should deal separately with the cases $a = 0$, $a \neq 0$.

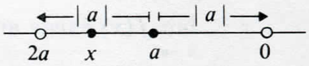
When $a = 0$, we see from (1) that

$$|f(x) - f(0)| = |x|\,|x| = |x|^2.$$

Hence, for given $\varepsilon > 0$, let us choose $\delta = \varepsilon^{\frac{1}{2}}$ It follows that if $|x| < \delta$ then $|x|^2 < \varepsilon$, and so $|f(x) - f(0)| < \varepsilon$; thus f is continuous at 0.



the case $a > 0$



the case $a < 0$

When $a \neq 0$, the argument has to be rather different. Loosely, we are trying to show that if $|x - a|$ is small, then so is $|f(x) - f(a)|$ (i.e. that if x is near to a then $f(x)$ is near to $f(a)$). We are interested only in the behaviour of f near a; there the term $|x + a|$ in (1) is more or less just $|2a|$, and so gives a (nearly) constant multiple of $|x - a|$ in (1). Pinning down what we mean by "near", we argue as follows.

If $|x - a| < |a|$, then

$$\begin{aligned} |x + a| &= |(x - a) + 2a| \\ &\leqslant |x - a| + 2|a| \qquad \text{(by the Triangle Inequality)} \\ &< 3|a|, \end{aligned}$$

so that

$$|f(x) - f(a)| < 3|a| \cdot |x - a|. \qquad (2)$$

To make $|f(x) - f(a)| < K\varepsilon$ for some fixed K, it is then enough to have firstly that $|x - a| < |a|$ (so that (2) holds) and secondly to ensure that the choice of δ guarantees that $3|a|\,|x - a| \leqslant K\varepsilon$ (for some K).

Hence, for given $\varepsilon > 0$, we choose $\delta = \min(|a|, \varepsilon)$. The above argument then shows that if $|x - a| < \delta$,

$$\begin{aligned} |f(x) - f(a)| &< 3|a|\delta \\ &\leqslant (3|a|)\varepsilon; \end{aligned}$$

hence, by Lemma 15, f is continuous at a.

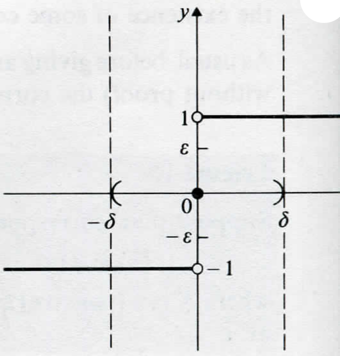
So f is continuous at each point of $\mathbf{R}$.

Example

Show that the function

$$f: x \longrightarrow \begin{cases} -1, & \text{if } x < 0, \\ 0, & \text{if } x = 0, \\ +1, & \text{if } x > 0, \end{cases}$$

is not continuous (i.e. is discontinuous) at 0.



Solution

As usual when trying to show that something does not happen, we shall use a contradiction argument.

Suppose that $f(x)$ is continuous at 0. Then, corresponding to each given $\varepsilon > 0$ there exists a $\delta > 0$ such that

$$|f(x) - f(0)| < \varepsilon \qquad \text{for all } x \text{ with } |x - 0| < \delta;$$

that is,

$$|f(x)| < \varepsilon \qquad \text{for all } x \text{ with } |x| < \delta.$$

But we notice that any small interval $(-\delta, \delta)$ about 0 contains points where f takes values $+1$ and -1, so if we choose a small ε to begin with, we are likely to obtain a contradiction.

So, let $\varepsilon = \frac{1}{2}$. Then there must exist some $\delta > 0$ such that

$$|f(x)| < \tfrac{1}{2} \qquad \text{for all } x \text{ with } |x| < \delta.$$

This last line is not true, since if we choose $x = \frac{1}{2}\delta$ we then obtain the inequality $1 = |f(\frac{1}{2}\delta)| < \frac{1}{2}$. This contradiction shows that f must be discontinuous at 0 after all.

Self-assessment Question

(a) Show that the function

$$f(x) = \begin{cases} x \sin\left(\dfrac{1}{x}\right), & \text{if } x \neq 0, \\ 0, & \text{if } x = 0, \end{cases}$$

is continuous at 0.

(b) Show that the function

$$f(x) = x^3$$

is continuous at 1.

(c) Show that the function

$$f(x) = \begin{cases} \sin\left(\dfrac{1}{x}\right), & \text{if } x \neq 0, \\ 0, & \text{if } x = 0, \end{cases}$$

is not continuous at 0.

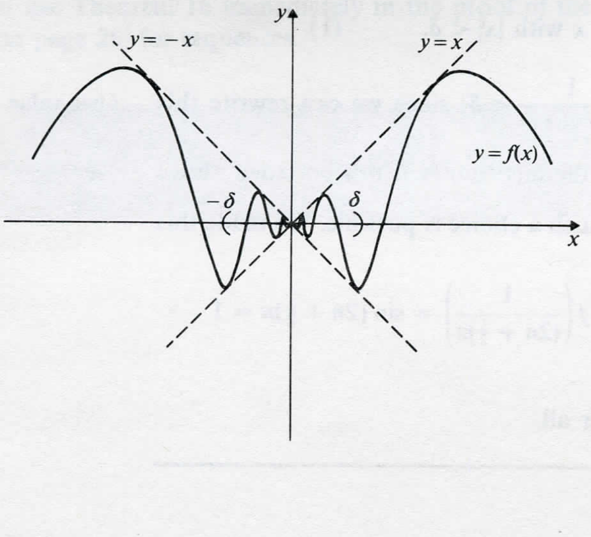
Solution

(a) Let $\varepsilon > 0$ be given.
Choose $\delta = \varepsilon$. Then, for all x with $|x - 0| < \delta$ we have that

$$\begin{aligned} |f(x) - f(0)| &= |f(x)| \\ &\leqslant |x| \\ &< \delta = \varepsilon; \end{aligned}$$

For $x \neq 0$ this holds because $|\sin x| \leqslant 1$ for all x, and for $x = 0$ it is obvious.

so that f is continuous at 0.



(b) Let $\varepsilon > 0$ be given.

We know that we should look at the difference

$$|f(x) - f(1)| = |x^3 - 1| = |x - 1| \cdot |x^2 + x + 1|.$$

For x close to 1, $|x - 1|$ will be small, and $x^2 + x + 1$ approximately equal to 3. So if we restrict our attention to $x \in (0, 2)$,

$$\begin{aligned} |x^2 + x + 1| &\leqslant |x|^2 + |x| + 1 \\ &\leqslant 2^2 + 2 + 1 = 7, \end{aligned}$$

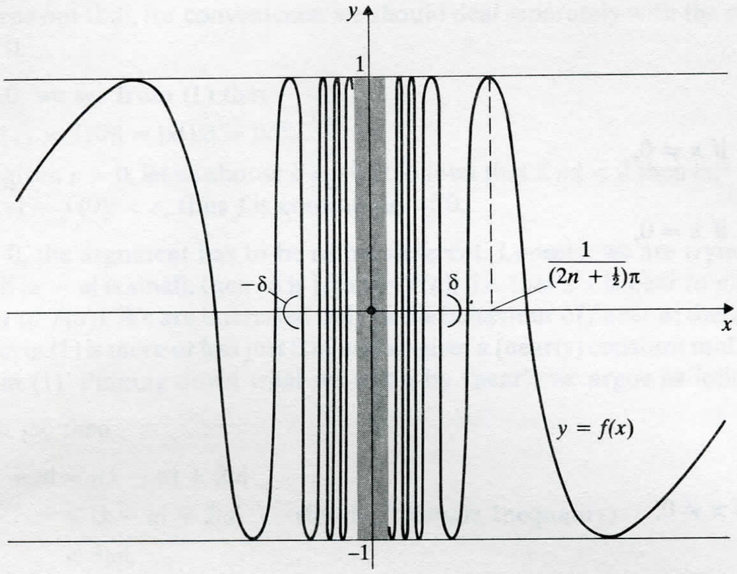
7 is certainly not approximately equal to 3; but all we need is *some* bound.

by the Triangle Inequality. If we also insist that $x \in (1 - \varepsilon, 1 + \varepsilon)$, we have $|x - 1| < \varepsilon$ in the other factor above. Combining these observations, we see that we can choose $\delta = \min(\varepsilon, 1)$, so that both the above bounds hold. Thus, for all x with $|x - 1| < \delta$,

$$|f(x) - f(1)| < 7\varepsilon.$$

The $K\varepsilon$ Lemma (Lemma 15) with $K = 7$ then shows that f is continuous at 1.

(c)



Our argument is very similar to that of the example on page 42; for whatever choice of δ we make, there are points in $(-\delta, \delta)$ where f takes both the values ± 1.

Suppose that f is continuous at 0. We shall seek a contradiction.

Choose $\varepsilon = \frac{1}{2}$. Then from the definition of continuity there must exist a corresponding value of $\delta > 0$ such that

$$|f(x) - f(0)| < \frac{1}{2} \quad \text{for all } x \text{ with } |x - 0| < \delta,$$

that is, that

$$|f(x)| < \frac{1}{2} \quad \text{for all } x \text{ with } |x| < \delta. \qquad (1)$$

Now choose an $n \in \mathbf{N}$ such that $\frac{1}{(2n + \frac{1}{2})\pi} < \delta$; since we can rewrite this inequality in the form $n > \frac{1}{2\pi\delta} - \frac{1}{4}$, the fact that $\mathbf{N}$ is not bounded above (Theorem 5 on page 16) shows that such a choice is possible. We made this choice of n since the facts that

f has value 1 at $\frac{1}{(2n + \frac{1}{2})\pi}$.

$$\frac{1}{(2n + \frac{1}{2})\pi} \in (-\delta, \delta) \quad \text{and} \quad f\left(\frac{1}{(2n + \frac{1}{2})\pi}\right) = \sin(2n + \tfrac{1}{2})\pi = 1$$

give us a contradiction with (1).

So f must be discontinuous at 0 after all.

3.4 Some Properties of Continuous Functions

In this section we are going to look at some of the different ways of combining continuous functions and then at some of the interesting (and occasionally surprising) properties of continuous functions on closed intervals.

Our first result will be particularly useful in the rest of the section; it is the analogue of Theorem 9 (for sequences).

Theorem 16

Let f be continuous at a. Then there exists a $\delta > 0$ such that f is bounded on the interval $(a - \delta, a + \delta)$.

Proof

We use $\varepsilon = 1$ in the definition of continuity. Since f is continuous at a, there exists a $\delta > 0$ such that

$$|f(x) - f(a)| < 1 \qquad \text{for all } x \text{ with } |x - a| < \delta.$$

Hence, for all $x \in (a - \delta, a + \delta)$, we have that

$$\begin{aligned} |f(x)| &= |(f(x) - f(a)) + f(a)| \qquad \text{(our standard trick!)} \\ &\leqslant |f(x) - f(a)| + |f(a)| \\ &< 1 + |f(a)|. \end{aligned}$$

It follows that f is bounded on $(a - \delta, a + \delta)$. ■

Self-assessment Question

Show that if f is continuous at a and $f(a) \neq 0$, then there exists a $\delta > 0$ such that $f(x)$ has the same sign at all points of $(a - \delta, a + \delta)$.
(Hint: Use $\varepsilon = \frac{1}{2}|f(a)|$ in the definition of continuity.)

Solution

Since $\frac{1}{2}|f(a)| > 0$ and f is continuous at a, there must exist some $\delta > 0$ such that

$$|f(x) - f(a)| < \tfrac{1}{2}|f(a)| \qquad \text{for all } x \text{ with } |x - a| < \delta.$$

Hence, for all $x \in (a - \delta, a + \delta)$ we have that

$$-\tfrac{1}{2}|f(a)| < f(x) - f(a) < \tfrac{1}{2}|f(a)|,$$

which we choose to rewrite in the more convenient form

$$f(a) - \tfrac{1}{2}|f(a)| < f(x) < f(a) + \tfrac{1}{2}|f(a)|.$$

If $f(a) > 0$ so that $|f(a)| = f(a)$, the left inequality gives us that $f(x) > \frac{1}{2}f(a) > 0$; and, if $f(a) < 0$ so that $|f(a)| = -f(a)$, the right inequality gives that $f(x) < \frac{1}{2}f(a) < 0$. In either case it follows that, for all $x \in (a - \delta, a + \delta)$, $f(x)$ has the same sign as $f(a)$.

We are able to use Theorem 16 immediately in the proof of the analogue of Theorem 10 (on page 26) for sequences.

Theorem 17

Let f, g be continuous at a, and let $k \in \mathbf{R}$. Then the following functions are continuous at a: (i) kf; (ii) $f + g$; (iii) fg.

Proof

(i) If $k = 0$, then for any given $\varepsilon > 0$ and for any choice of $\delta > 0$ we have that

$$|kf(x) - kf(a)| < \varepsilon \qquad \text{for all } x \text{ with } |x - a| < \delta;$$

so that kf is continuous at a.

So now suppose that $k \neq 0$. Let $\varepsilon > 0$ be given. Since f is continuous at a, there exists $\delta > 0$ such that

$$|f(x) - f(a)| < \varepsilon \qquad \text{for all } x \text{ with } |x - a| < \delta.$$

It follows that

$$|kf(x) - kf(a)| < |k|\varepsilon \qquad \text{for all } x \text{ with } |x - a| < \delta;$$

and so, by "the $K\varepsilon$ Lemma", with $K = |k| > 0$, kf is continuous at a.

(ii) This proof is left to an SAQ on page below.

(iii) Choose any $\varepsilon > 0$.

Since f, g are continuous at a, there exist $\delta_1 > 0$ and $\delta_2 > 0$ such that

$$|f(x) - f(a)| < \varepsilon \qquad \text{for all } x \text{ with } |x - a| < \delta_1, \text{ and} \qquad (1)$$

$$|g(x) - g(a)| < \varepsilon \qquad \text{for all } x \text{ with } |x - a| < \delta_2. \qquad (2)$$

Now, we are interested in the expression $|f(x)g(x) - f(a)g(a)|$ for x close to a. In order to get into a position where we can use the information contained in (1) and (2), we see that a regrouping of the expression in the form

$$|f(x)\{g(x) - g(a)\} + g(a)\{f(x) - f(a)\}| \qquad (3)$$

will certainly help. For a start we see that the expression (3) is, by the Triangle Inequality, less than or equal to

$$|f(x)| \cdot |g(x) - g(a)| + |g(a)| \cdot |f(x) - f(a)|. \qquad (4)$$

Now, (1) and (2) tell us that two of these four terms are small if x is near a; also, $|g(a)|$ is a constant, and by Theorem 16 we know that $|f(x)|$ is at least bounded near a (so that its product with $|g(x) - g(a)|$ will be small).

So suppose that, by Theorem 16, we choose $\delta_3 > 0$ such that

$$|f(x)| < 1 + |f(a)| \qquad \text{for all } x \text{ with } |x - a| < \delta_3. \qquad (5)$$

Now let us put together all the information we have.

Choose $\delta = \min(\delta_1, \delta_2, \delta_3)$, so that for those x for which $|x - a| < \delta$, all the inequalities in (1), (2) and (5) are valid. Hence for x with $|x - a| < \delta$,

$$\begin{aligned} |f(x)g(x) - f(a)g(a)| &\leqslant |f(x)| \cdot |g(x) - g(a)| \\ &\quad + |g(a)| \cdot |f(a) - f(a)| \quad \text{(by (3) and (4))} \\ &< (1 + |f(a)|)\varepsilon + |g(a)|\varepsilon \\ &= (1 + |f(a)| + |g(a)|)\varepsilon. \end{aligned}$$

Hence, by the $K\varepsilon$ Lemma with $K = 1 + |f(a)| + |g(a)|$, it follows that fg is continuous at a. ■

Self-assessment Question

Prove part (ii) of Theorem 17.

Solution

Let $\varepsilon > 0$. Since f is continuous at a, there exists a $\delta_1 > 0$ such that

$$|f(x) - f(a)| < \varepsilon \qquad \text{for all } x \text{ with } |x - a| < \delta_1;$$

and since g is continuous at a, there exists a $\delta_2 > 0$ such that

$$|g(x) - g(a)| < \varepsilon \qquad \text{for all } x \text{ with } |x - a| < \delta_2.$$

So if we choose $\delta = \min(\delta_1, \delta_2)$, both inequalities will hold; in other words, for all x with $|x - a| < \delta$,

$$|f(x) - f(a)| < \varepsilon \quad \text{and} \quad |g(x) - g(a)| < \varepsilon.$$

We know that we have to look at the expression $|(f+g)(x) - (f+g)(a)|$, and notice that we can regroup these terms as $|\{f(x) - f(a)\} + \{g(x) - g(a)\}|$. This suggests at once that the Triangle Inequality will help to relate this difference to the expressions we have just been looking at. So, for all x with $|x - a| < \delta$, we see that

$$\begin{aligned} |(f+g)(x) - (f+g)(a)| &= |\{f(x) - f(a)\} + \{g(x) - g(a)\}| \\ &\leqslant |f(x) - f(a)| + |g(x) - g(a)| \\ &< \varepsilon + \varepsilon = 2\varepsilon. \end{aligned}$$

The $K\varepsilon$ Lemma, with $K = 2$, shows that $f + g$ is continuous at a.

There are many other results for continuous functions that we do not specifically mention here. (For example, if f, g are continuous then so are $|f|$, $\frac{1}{f}$ and $f \circ g$; and if a function f is continuous and one-one on an interval I, then its inverse function f^{-1} is also continuous.)

We must have $f(x) \neq 0$ for $\frac{1}{f}$ to be continuous.

In *Unit 2*, we shall consider continuous functions whose domain and codomain are both **C**, and shall see there how many of our results for real functions can be generalized to complex functions.

We end this section by reminding you of two important consequences of continuity that we shall often require. First, some more terminology.

Definition

We say that the function $f\colon [a,b] \longrightarrow \mathbf{R}$ is **continuous on the closed interval** $[a,b]$ if the associated function $g\colon \mathbf{R} \longrightarrow \mathbf{R}$ defined by

$$g(x) = \begin{cases} f(x), & \text{if } a \leqslant x \leqslant b, \\ f(a), & \text{if } x < a, \\ f(b), & \text{if } x > b, \end{cases}$$

is continuous at each point of $[a,b]$.

Do not spend too much time on this definition.

The point of the definition is that it enables us to check for continuity of f at a and at b, since otherwise we do not have open intervals containing a and b in the domain of f, and so cannot apply the definition of continuity on page 41 at the points a and b.

Theorem 18 (The Boundedness Property)

Let f be a continuous function on the *closed* interval $[a,b]$. Then

(i) f is bounded on $[a,b]$; and

(ii) f attains its upper and lower bounds on $[a,b]$.

Theorem 19 (The Intermediate Value Property)

Let f be a continuous function on the *closed* interval $[a,b]$, and let d be some real number lying between the values $f(a)$ and $f(b)$. Then there exists some point c between a and b where $f(c) = d$.

3.5 Uniform Continuity

In various places in *M332*, we shall find that continuous functions are not quite sufficiently well-behaved for our purposes. So we now look at continuous functions that satisfy one additional condition; these are called *uniformly continuous* functions.

In the definition on page 41 of continuity of a function f at a point a of some interval I, we allowed our choice of δ to depend on the point a (as well as on f, I and, of course, ε). If, however, we require that the same value of δ will work for *all* points a in I, we end up with

Definition

Let f be defined on an interval I. Then f is **uniformly continuous** on I if, corresponding to every given $\varepsilon > 0$, there exists some $\delta > 0$ (whose value does not depend on a) such that, for all a and x in I with $|x - a| < \delta$, then $|f(x) - f(a)| < \varepsilon$.

It is obvious from the definition that a uniformly continuous function on an interval I must be continuous on I. However, a function that is continuous on an interval I is not necessarily uniformly continuous; for example, if we look closely at the function $f(x) = \dfrac{1}{x}$ on the open interval $(0, 1)$, we see that corresponding to any given $\varepsilon > 0$, there is no value of δ such that the condition in the above definition is satisfied.

We ignore the justification of o
assertions here.

Nevertheless, there is one special (and important) situation where if a function is continuous it is necessarily uniformly continuous.

Theorem 20

Let f be continuous on the closed interval $[a, b]$. Then f is uniformly continuous on $[a, b]$.

This is a quite deep result of real analysis, and we omit its proof.

You will meet uniform continuity again on page 95 of *Unit 2*, and it will play a vital role in our work in *Unit 5* and later.

4 DIFFERENTIATION

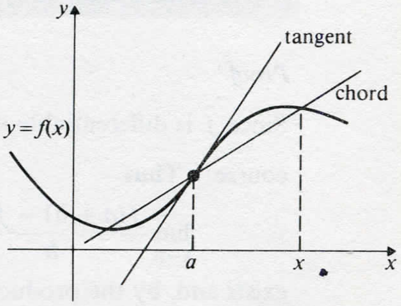
4.1 The Derivative

In this section we restate in *M332* terminology the definition of the derivative, and look at ways of combining differentiable functions. Our starting point is

Definition 1

> The function f defined on some open interval I is said to be **differentiable at a point** $a \in I$ if $\lim_{x \to a} \frac{f(x) - f(a)}{x - a}$ exists.

When f is differentiable, the value of the limit is called the **derivative of** f **at** a, and is denoted by $f'(a)$. The function $f' : x \longrightarrow f'(x)$ is called the **derivative** or **derived function** of f.



We can interpret the derivative, $f'(a)$, geometrically in the following way. The quotient $\frac{f(x) - f(a)}{x - a}$ represents the slope of the chord joining the points $(a, f(a))$ and $(x, f(x))$ on the curve $y = f(x)$. As the tangent to the curve at the point $(a, f(a))$ is the limiting position of the chord as x approaches a, the derivative $f'(a)$ is just the *slope* of the tangent to the curve at $(a, f(a))$.

For many purposes, it is more convenient to rephrase the definition of $f'(a)$ in the following two equivalent forms.

Definition 2

> If f is **differentiable** at a,
>
> $$f'(a) = \lim_{h \to 0} \frac{f(a+h) - f(a)}{h}.$$

This is just Definition 1 with x replaced by $a + h$, so that $x \to a$ is just the same condition as $h \to 0$.

Definition 3

> f is **differentiable** at a with **derivative** $f'(a)$ if the following condition is satisfied:
>
> corresponding to each given $\varepsilon > 0$ there exists a $\delta > 0$ such that
>
> $$\left| \frac{f(x) - f(a)}{x - a} - f'(a) \right| < \varepsilon \qquad \text{for all } x \text{ with } 0 < |x - a| < \delta$$
>
> (or, equivalently, such that
>
> $$\left| \frac{f(a+h) - f(a)}{h} - f'(a) \right| < \varepsilon \qquad \text{for all } h \text{ with } 0 < |h| < \delta).$$

This is simply Definitions 1 and 2 with the limit statements expressed in terms of ε and δ.

We shall use analogues of each version of the definition of $f'(a)$ at various points throughout *M332*; see, for a sneak preview, how on page 78 of *Unit 5* we show that some expression is the derivative of f at z by looking at the difference

$$\left|\frac{f(z+h)-f(z)}{h} - \text{the expression}\right|$$

for values of h near to 0.

Without stating it explicitly, we remark that the analogue of our early "$K\varepsilon$ Lemmas" applies equally to the ε in Definition 3.

Since we would expect the graph $y = f(x)$ of a function f to have a tangent only if the graph had "no breaks" in it, the following result is no surprise.

Theorem 21

If f is differentiable at a, then it is continuous at a.

Proof

Since f is differentiable at a, $\lim\limits_{h\to 0} \dfrac{f(a+h)-f(a)}{h}$ exists; and so does $\lim\limits_{h\to 0} h$ (of course!). Thus

$$\lim_{h\to 0} \frac{f(a+h)-f(a)}{h}\cdot h = \lim_{h\to 0} \{f(a+h)-f(a)\}$$

exists and, by the product rule for limits, equals

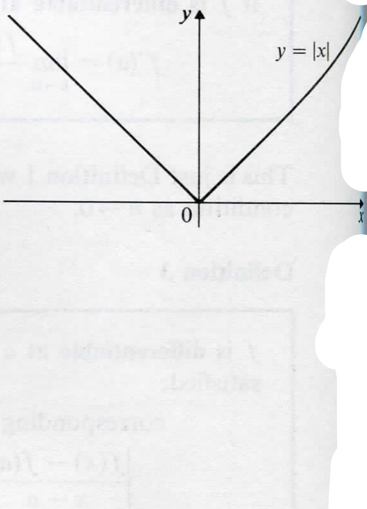
$$\begin{aligned}\lim_{h\to 0} \frac{f(a+h)-f(a)}{h}\cdot\lim_{h\to 0} h &= f'(a)\cdot 0\\ &= 0.\end{aligned}$$

So $\lim\limits_{h\to 0} f(a+h)$ exists and equals $f(a)$; therefore f is continuous at a. ■

Notice, however, that the converse to Theorem 21 is false; that is, f can be continuous at a point a but not differentiable at a.

Example

Let $f(x) = |x|$. It is not hard to verify that f is continuous at 0 but is not differentiable at 0.



We now end this section by looking at the derivative of a specific function $f(x)$ and at some of the (standard) ways of combining the derivatives of different functions.

Example

Find the derivative of the function $f(x) = x^3$ at a point a, directly from the definition of derivative.

Solution

First let us use Definition 3 to verify that $f'(a) = 3a^2$.

Let $\varepsilon > 0$ be given. We know that we should look at the expression

$$\begin{aligned}\left|\frac{f(x)-f(a)}{x-a} - 3a^2\right| &= \left|\frac{x^3-a^3}{x-a} - 3a^2\right|\\ &= |x^2+ax-2a^2|\\ &= |(x^2-a^2)+a(x-a)| && \text{(regrouping)}\\ &\leqslant |x^2-a^2| + |a|\,|x-a| && \text{(by the Triangle Inequality)}\\ &= |x-a|\,|x+a| + |a|\,|x-a|\\ &\leqslant |x-a|\,(|x|+2|a|) && \text{(by the Triangle Inequality).}\end{aligned}$$

So choose $\delta = \min(1, \varepsilon)$. Then for all x with $0 < |x - a| < \delta$, we have $|x| < |a| + \delta \leqslant |a| + 1$; and from our above discussion it follows that

$$\left|\frac{f(x) - f(a)}{x - a} - 3a^2\right| < \varepsilon(3|a| + 1).$$

It follows (using "the $K\varepsilon$ Lemma") that $f'(a)$ exists and equals $3a^2$.

Quite often in *M332* we shall show that some function f is differentiable using the method we have just outlined. Equally often, though, we shall use the limit approach of Definition 2 together with the rules for combining sums and products of limits. Let us see now how we would use this approach in the present example.

We know that we should look at the quotient

$$\begin{aligned}\frac{f(a+h) - f(a)}{h} &= \frac{(a+h)^3 - a^3}{h} \\ &= \frac{a^3 + 3a^2h + 3ah^2 + h^3 - a^3}{h} \\ &= 3a^2 + 3ah + h^2.\end{aligned}$$

If we let $h \to 0$ in this identity,

$$\begin{aligned}\lim_{h\to 0} \frac{f(a+h) - f(a)}{h} &= \lim_{h\to 0} (3a^2 + 3ah + h^2) \\ &= \lim_{h\to 0} (3a^2) + \lim_{h\to 0} (3ah) + \lim_{h\to 0} (h^2) \\ &= 3a^2,\end{aligned}$$

it follows that $f'(a)$ exists and equals $3a^2$.

Sometimes one definition is simpler to apply than another, and in any given situation we choose the most convenient one. (Here obviously Definition 2 was the simplest one.)

Just as we were able to verify the standard rules for combining sequences or continuous functions, we could use our "ε-δ" or "h" definitions of differentiation to verify

Theorem 22

Let f and g be differentiable functions, and let $k \in \mathbf{R}$. Then the following functions are also differentiable: kf; $f + g$; fg; $\frac{1}{f}$ (at points where $f \neq 0$); $f \circ g$.

We ask you to look at one of these in an SAQ below; and you will see the analogous result for complex functions in Section 3.1 of *Unit 3*.

Self-assessment Question

Prove that if f and g are differentiable at a, then so is $f + g$.

Solution

We use Definition 3, in terms of h.

For $h \neq 0$,

$$\begin{aligned}\frac{(f+g)(a+h) - (f+g)(a)}{h} &= \frac{\{f(a+h) - f(a)\} + \{g(a+h) - g(a)\}}{h} \\ &= \frac{f(a+h) - f(a)}{h} + \frac{g(a+h) - g(a)}{h}.\end{aligned}$$

Now let $h \to 0$ and use the fact that f and g are differentiable at a; it follows that

$$\lim_{h\to 0} \frac{(f+g)(a+h)-(f+g)(a)}{h} = \lim_{h\to 0} \frac{f(a+h)-f(a)}{h} + \lim_{h\to 0} \frac{g(a+h)-g(a)}{h}$$
$$= f'(a) + g'(a).$$

In other words, $f + g$ is differentiable at a, and its derivative is $f'(a) + g'(a)$.

4.2 Properties of the Derivative

For use later in the course, it is convenient to remind you in this section (without any proofs at all) of some of the most useful consequences of differentiability.

Theorem 23 (Rolle's Theorem)

Let $f(x)$ be continuous on the closed interval $[a, b]$, differentiable on the open interval (a, b), and have $f(a) = f(b)$. Then there exists some point $c \in (a, b)$ such that

$$f'(c) = 0.$$

Theorem 24 (Mean Value Theorem)

Let $f(x)$ be continuous on the closed interval $[a, b]$ and differentiable on the open interval (a, b). Then there exists some point $c \in (a, b)$ such that

$$f'(c) = \frac{f(b) - f(a)}{b - a}.$$

Theorem 25 (L'Hôpital's Rule)

Suppose that the functions f, g are differentiable at each point of an open interval $(a - \delta, a + \delta)$, and that $f(a) = g(a) = 0$. Then

$$\lim_{x\to a} \frac{f(x)}{g(x)} \text{ exists if } \lim_{x\to a} \frac{f'(x)}{g'(x)} \text{ exists;}$$

when this happens, the two limits are equal. In particular, if $g'(a) \neq 0$,

$$\lim_{x\to a} \frac{f(x)}{g(x)} = \frac{f'(a)}{g'(a)}.$$

You will see l'Hôpital's Rule again on page147 of *Unit* 6, and then fairly regularly in *Unit 8* (see page 15, for example).

Self-assessment Question

Use l'Hôpital's Rule to determine $\lim_{x\to 0} \frac{1 - \cos x}{x^2}$.

Solution

Here we apply Theorem 25 with $a = 0$, $f(x) = 1 - \cos x$ and $g(x) = x^2$. f and g are differentiable at 0, and $f(0) = g(0) = 0$. Hence

$$\lim_{x\to 0} \frac{1 - \cos x}{x^2} \text{ exists if} \qquad (1)$$

$$\lim_{x\to 0} \frac{f'(x)}{g'(x)} = \lim_{x\to 0} \frac{\sin x}{2x} \text{ exists;} \qquad (2)$$

in this case the limits would be equal. To see whether this last limit exists, we have to use Theorem 25 again. Both the numerator and denominator are differentiable at 0 and vanish at 0. Hence (2) exists if

$$\lim_{x \to 0} \frac{\cos x}{2} \qquad (3)$$

exists, and if so then (2) and (3) are equal. In fact, (3) does exist and equals $\frac{1}{2}$; so, going back through the argument, (2) and thus (1) both exist and equal $\frac{1}{2}$.

Theorems 23, 24 and 25 were proved in *M203, Analysis Block B, Unit 1, Differentiation.* Theorem 26 (below) was discussed in *M203, Analysis Block B, Unit 3, Power Series.*

Theorem 26 (Taylor's Theorem)

Let the function f have as domain $I = (a - \delta, a + \delta)$ for some $\delta > 0$. Suppose that f can be differentiated $(n + 1)$ times at each point of I. Then for any $x \in I$ there exists a point c between a and x such that

$$f(x) = f(a) + f'(a)(x - a) + \frac{f''(a)}{2!}(x - a)^2 + \ldots + \frac{f^{(n)}(a)}{n!}(x - a)^n + R_n$$

where the "remainder term" $R_n = \dfrac{f^{(n+1)}(c)}{(n+1)!}(x - a)^{n+1}$.

As you will have seen in your previous work in analysis, we can express an infinitely-differentiable function f in the form

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!}(x - a)^n, \qquad x \in (a - \delta, a + \delta) \qquad (4)$$

only if we are sure that $\lim_{n \to \infty} R_n = 0$. When this happens, (4) is called the **Taylor Series for f at a.** For example, for the function

$$f(x) = \begin{cases} 0, & \text{if } x = 0, \\ e^{-1/x^2}, & \text{if } x \neq 0, \end{cases}$$

we can show (with some work) that $f^{(n)}(0) = 0$ for each n; so that for this particular function it is *not* true that

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(0)}{n!} x^n.$$

As you will see in *Unit 6* (page 130, for example), the behaviour of Taylor Series in **C** is actually much simpler than that of Taylor Series in **R**.

Remark

Suppose that f is infinitely differentiable in the open interval $(a - \delta, a + \delta)$ for some $\delta > 0$ and that, for some real numbers a_n, $f(x)$ can be expressed in the form

$$f(x) = \sum_{n=0}^{\infty} a_n(x - a)^n, \qquad x \in (a - \delta, a + \delta). \qquad (5)$$

By carefully checking through the proof of Theorem 26, we can show that then, necessarily, $a_n = \dfrac{f^{(n)}(a)}{n!}$. In other words, if a function has a Taylor expansion as in (5), then the Taylor coefficients are *unique.*

We shall capitalize on this uniqueness property in the following way. Theorem 26 tells us how to determine the Taylor coefficients a_n by repeated differentiation. We shall meet several other methods of determining the a_n (for example, by using partial fractions on page 60); their efficacy depends on the fact that *any* method of determining coefficients a_n in an expansion of the form (5) necessarily gives the same values of the a_n as does Theorem 26.

Example

Determine the Taylor Series for $f(x) = x^3$ at 1.

Solution

$$\begin{aligned} \text{Here } f(x) &= x^3 && \text{so that } f(1) &&= 1;\\ f'(x) &= 3x^2 && \text{so that } f'(1) &&= 3;\\ f''(x) &= 6x && \text{so that } f''(1) &&= 6;\\ f'''(x) &= 6 && \text{so that } f'''(1) &&= 6; \end{aligned}$$

and then $f^{(4)}(x) = 0$, so that $f^{(n)}(x) = 0$ for all $n \geqslant 4$. In particular, if we apply Theorem 26 with $a = 1$ and $n = 3$, it follows that

$$\begin{aligned} x^3 &= f(x)\\ &= f(1) + f'(1)(x-1) + \frac{f''(1)}{2!}(x-1)^2 + \frac{f'''(1)}{3!}(x-1)^3\\ &\quad + \frac{f^{(4)}(c)}{4!}(x-1)^4 \qquad \text{(for some point } c \text{ between 0 and 1)}\\ &= 1 + 3(x-1) + 3(x-1)^2 + (x-1)^3 \end{aligned}$$

(which is what you would expect from expanding $(1 + (x-1))^3$).

Self-assessment Question

Determine the Taylor Series for $f(x) = e^x$ at 0.

Solution

By mathematical induction, we can show that $f^{(n)}(x) = e^x$ for each positive integer n. Hence $f^{(n)}(0) = 1$ for all $n \geqslant 1$.

In this case, the remainder term in Theorem 26 is

$$R_n = \frac{f^{(n+1)}(c)}{(n+1)!}x^{n+1} = \frac{x^{n+1}e^c}{(n+1)!},$$

which tends to 0 as $n \to \infty$ in view of the Remark on page 26. It follows that the Taylor Series for f at 0 is

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!} \qquad \text{for all } x \in \mathbf{R}.$$

Self-assessment Question

Determine the Taylor Series for $g(x) = e^{x^2}$ at 0.

Solution

Obviously, it is not sensible to use Theorem 26 directly here, as the formula for the nth derivative of g is pretty inconvenient. However, if we simply replace x by x^2 in the Taylor expansion in the previous SAQ, we see that

$$g(x) = f(x^2) = \sum_{n=0}^{\infty} \frac{x^{2n}}{n!} \qquad (x \in \mathbf{R}).$$

5 SERIES AND POWER SERIES

Here we remind you of the main properties of series $\sum a_n$ and power series $\sum a_n(x-a)^n$ for real a_n, x and a. We shall meet their analogues in **C** later in *M332*, from *Unit 6*, page 106 onwards. As mentioned in the Introduction, our account is exceedingly condensed, so if you feel that you have forgotten any of the results we mention, you should look back to your previous course in analysis or to **Spivak.**

5.1 Convergence and Absolute Convergence of Series

Definition

An infinite series $\sum_{n=1}^{\infty} a_n$ (of real numbers a_n) is said to **converge** to the **sum** s if the sequence $\{s_n\}$, where $s_n = \sum_{k=1}^{n} a_k = a_1 + a_2 + \ldots + a_n$, converges to the limit s. If a series does not converge it is said to **diverge.**

Of course, it is all very well to make such definitions; but they are of real value only if, given a series, we have a criterion which enables us to determine whether the series is convergent or divergent. Our first result is a step in the right direction.

Theorem 27

If an infinite series $\sum_{n=1}^{\infty} a_n$ is convergent, then $a_n \to 0$ as $n \to \infty$.

Thus, for example, we can conclude without any real effort that the series $\sum_{n=1}^{\infty} (-1)^n n$ must be divergent.

It is very tempting to hope that the converse to Theorem 27 would be true; that is, that if $a_n \to 0$ as $n \to \infty$, then the series $\sum_{n=1}^{\infty} a_n$ is necessarily convergent. Unfortunately, this is not so, as we see below.

Example (*The Harmonic Series*)

Show that the series $\sum_{n=1}^{\infty} \frac{1}{n}$ is divergent.

Solution

Let us look at this claim, using a contradiction argument. Suppose that the series $\sum_{n=1}^{\infty} \frac{1}{n}$ were convergent, to some sum s. Then by applying Theorem 9 (on page 25) to the sequence $s_n = \sum_{k=1}^{n} \frac{1}{k}$, we see that some M exists such that

$$s_n \leqslant M \qquad \text{for all } n \in \mathbf{N}. \qquad (1)$$

Now we examine carefully the sequence $\{s_n\}$ of partial sums for the harmonic series. Obviously, $s_1 = 1$, $s_2 = 1 + \frac{1}{2}$. Ignoring s_3 (for reasons that will become clear), we then notice that

$$\begin{aligned} s_4 &= 1 + \tfrac{1}{2} + (\tfrac{1}{3} + \tfrac{1}{4}) \\ &> 1 + \tfrac{1}{2} + 2 \times (\tfrac{1}{4}), \end{aligned}$$

which we prefer to rewrite in the form

$$s_{2^2} > 1 + 2 \times (\tfrac{1}{2}).$$

The next s_n of interest to us is then s_8, for which

$$\begin{aligned} s_{2^3} &= s_8 \\ &= 1 + \tfrac{1}{2} + (\tfrac{1}{3} + \tfrac{1}{4}) + (\tfrac{1}{5} + \tfrac{1}{6} + \tfrac{1}{7} + \tfrac{1}{8}) \\ &> 1 + \tfrac{1}{2} + 2 \times (\tfrac{1}{4}) + 4 \times (\tfrac{1}{8}) \\ &= 1 + 3 \times (\tfrac{1}{2}). \end{aligned}$$

Indeed, using the same trick of grouping terms appropriately and using mathematical induction, we can prove that

$$s_{2^m} > 1 + \tfrac{1}{2}m. \qquad (2)$$

However, since **N** is not bounded above, it follows that there exists some positive integer m such that $1 + \frac{1}{2}m > M$ and so

$$s_{2^m} > M. \qquad (3)$$

As (1) and (3) are inconsistent, we have found a contradiction. We conclude that the harmonic series is divergent.

The harmonic series is a special case of the more general

Example

The series $\sum_{n=1}^{\infty} \frac{1}{n^p}$ is convergent if $p > 1$ and divergent if $p \leqslant 1$.

We quote this without proof.

The two most useful tests for convergence are given in

Theorem 28

(a) (Comparison Test) Suppose that $|a_n| \leqslant Kb_n$ where K is a constant (independent of n) and the series $\sum_{n=1}^{\infty} b_n$ is convergent. Then $\sum_{n=1}^{\infty} a_n$ is convergent.

(b) (Ratio Test) If $a_n > 0$ for all n and $\frac{a_{n+1}}{a_n} \to r$ as $n \to \infty$, then the series $\sum_{n=1}^{\infty} a_n$ converges if $r < 1$ and diverges if $r > 1$. (No conclusion may be drawn if $r = 1$; that is, the series may diverge or converge if $r = 1$.)

Self-assessment Question

Prove that, for each $n \in \mathbf{N}$, $\frac{n+1}{n^3 + 2n^2} < \frac{1}{n^2}$. Hence determine whether or not the series $\sum_{n=1}^{\infty} (-1)^n \frac{n+1}{n^3 + 2n^2}$ converges.

Solution

For each $n \in \mathbf{N}$,

$$\frac{n+1}{n^3 + 2n^2} < \frac{1}{n^2} \iff n^3 + n^2 < n^3 + 2n^2$$

$$\iff n^2 > 0,$$

$\iff$ means "if and only if".

and this inequality certainly holds.

Hence, since the series $\sum \frac{1}{n^2}$ converges, we can apply the Comparison Test with $a_n = \frac{n+1}{n^3 + 2n^2}$, $K = 1$, and $b_n = \frac{1}{n^2}$, to conclude that the given series converges.

This is the case $p = 2$ of the above example.

In our work on Taylor Series in *Unit 6*, we shall see that most of the convergent series that arise in Complex Analysis are not just convergent but are also absolutely convergent.

Definition

An infinite series $\sum_{n=1}^{\infty} a_n$ is said to be **absolutely convergent** if the corresponding series $\sum_{n=1}^{\infty} |a_n|$ is convergent.

Using Theorem 28(a), we can check that, *if a series* $\sum_{n=1}^{\infty} a_n$ *is absolutely convergent, it follows that it is convergent.* However, the converse result is false; for example, the series $\sum_{n=1}^{\infty} (-1)^n \frac{1}{n}$ is convergent, although the corresponding series of moduli, $\sum_{n=1}^{\infty} \frac{1}{n}$, is divergent.

This series is used often.

This is the harmonic series.

Self-assessment Question

Determine whether the series $\sum a_n$, where

$$a_n = \frac{(-1)^n}{n^n},$$

is (a) absolutely convergent, (b) convergent.

(Hint: use the fact that for each $\alpha, \beta \in \mathbf{R}$,

$$\lim_{n \to \infty} \left(1 + \frac{\alpha}{n}\right)^{\beta n} = e^{\alpha\beta}.)$$

This is a "standard" limit in real analysis.

Solution

(a) The series is absolutely convergent if the series $\sum b_n$, where $b_n = |a_n| = \frac{1}{n^n}$, is convergent. Now

$$\begin{aligned}\frac{b_{n+1}}{b_n} &= \frac{n^n}{(n+1)^{n+1}}\\ &= \frac{1}{n+1} \cdot \left(1 + \frac{1}{n}\right)^{-n} \to 0 \cdot e^{-1} = 0 \quad \text{as } n \to \infty;\end{aligned}$$

so that, by Theorem 28(b), $\sum b_n$ is convergent. Hence $\sum a_n$ is absolutely convergent.

(b) Since $\sum a_n$ is absolutely convergent, it is also convergent.

5.2 Power Series

In Complex Analysis, probably the most important type of infinite series is a power series.

Definition

A **power series** is an infinite series of the form $\sum_{n=0}^{\infty} a_n (x - a)^n$, where $a_n, x, a \in \mathbf{R}$.

Such series are of real interest because of the following crucial property.

Theorem 29

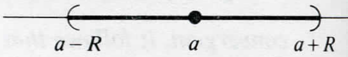
Corresponding to each power series $\sum_{n=0}^{\infty} a_n(x-a)^n$, there corresponds some real number R, $0 \leqslant R \leqslant \infty$, with the following properties:

(a) the power series converges, and converges absolutely, if $|x-a| < R$;

(b) the power series diverges if $|x-a| > R$.

$R = 0$ means that the series converg
only if $x - a = 0$; $R = \infty$ means t
it converges (absolutely) for all x.

(The theorem gives no information about the points x where $x - a = \pm R$; the series may converge or diverge at either.) The number R is called the **radius of convergence** of the power series, and $(a-R, a+R)$ the **interval of convergence.** The following result turns out to be one useful way of determining power series.



Theorem 30

Let R be the radius of convergence of the power series $\sum_{n=0}^{\infty} a_n(x-a)^n$, where $R > 0$; and, for each $x \in (a-R, a+R)$, let $f(x)$ denote the sum of the series. Then (a) f is differentiable in $(a-R, a+R)$, and

$$f'(x) = \sum_{n=1}^{\infty} n a_n (x-a)^{n-1};$$

(b) f is integrable in $(a-R, a+R)$, and

$$\int_a^x f(t)\,dt = \sum_{n=0}^{\infty} a_n \frac{(x-a)^{n+1}}{n+1} \quad \text{for } x \in (a-R, a+R).$$

So $\sum_{n=0}^{\infty} a_n(x-a)^n$ is the Taylor Ser
for f at a.

Remark

We may express the results of Theorem 30 non-technically in the following form: inside its interval of convergence, we may differentiate or integrate a power series *term-by-term*, and we get the "correct" answer.

Example (*The Binomial Series*)

For each $\alpha \in \mathbf{R}$,

$$(1+x)^\alpha = 1 + \alpha x + \frac{\alpha(\alpha-1)}{2!}x^2 + \ldots + \frac{\alpha(\alpha-1)\ldots(\alpha-n+1)}{n!}x^n + \ldots;$$

and this series is convergent in the interval $|x| < 1$.

Example (*The Geometric Series*)

For each x with $|x| < 1$,

$$\sum_{n=0}^{\infty} x^n = \frac{1}{1-x}.$$

This series diverges if $|x| \geqslant 1$.

Examples

For each $x \in \mathbf{R}$, we have

$$\sin x = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!} = x - \frac{x^3}{3!} + \frac{x^5}{5!} \ldots;$$

$$\cos x = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n}}{(2n)!} = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \ldots;$$

$$\exp x = \sum_{n=0}^{\infty} \frac{x^n}{n!} = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \frac{x^3}{3!} + \ldots.$$

Finally, for each $x \in (-1, 1)$,

$$\log_e(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \ldots = \sum_{n=1}^{\infty} (-1)^{n+1} \frac{x^n}{n};$$

this series diverges if $|x| > 1$.

Self-assessment Question

Find simple expressions for the sums of the series

(a) $\sum_{n=0}^{\infty} nx^n$; (b) $\sum_{n=1}^{\infty} \dfrac{x^n}{n(n+1)}$.

Solution

(a) Let $f(x)$ be the sum of the series. The expression for $f(x)$ looks very much like the derivative of the series for $\dfrac{1}{1-x}$, the Geometric Series. If we differentiate that series term-by-term, we obtain

$$\sum_{n=1}^{\infty} nx^{n-1} = \frac{1}{(1-x)^2}, \qquad \text{for } x \in (-1, 1); \qquad (*)$$

Notice that there is no term for $n = 0$. $(*)$ is used on page 62.

multiplying both sides by x, we deduce that

$$\sum_{n=1}^{\infty} nx^n = \frac{x}{(1-x)^2},$$

or, in other words, that

$$f(x) = \frac{x}{(1-x)^2}, \qquad \text{for } x \in (-1, 1).$$

For $|x| \geqslant 1$ the series is divergent.

(b) Let $g(x)$ be the sum of the given series. The expression for $g(x)$ looks rather like the integrated form of the right-hand side of the identity (for $x \in (-1, 1)$)

$$-\log_e (1-x) = x + \frac{x^2}{2} + \frac{x^3}{3} + \ldots = \sum_{n=1}^{\infty} \frac{x^n}{n}.$$

This is obtained from the series for $\log_e(1+x)$ in the previous example by putting $-x$ for x.

So, integrating both sides from 0 to x, we get that

$$\begin{aligned}
\sum_{n=1}^{\infty} \frac{x^{n+1}}{n(n+1)} &= -\int_0^x \log_e (1-t)\,dt \\
&= -\int_0^x 1 \cdot \log_e (1-t)\,dt \\
&= -\Big[t \log_e (1-t) \Big]_0^x - \int_0^x \frac{t}{1-t}\,dt \\
&\quad \text{(using integration by parts)} \\
&= -x \log_e (1-x) + \int_0^x \Big(1 - \frac{1}{1-t}\Big)\,dt \\
&= -x \log_e (1-x) + x + \log_e (1-x).
\end{aligned}$$

We use t as the "dummy variable", to avoid x having two meanings.

Here we use the fact that $\dfrac{t}{1-t} = 1 - \dfrac{1}{1-t}$.

Dividing both sides by x, it follows that

$$g(x) = -\log_e (1-x) + 1 + \frac{1}{x} \log_e (1-x) \qquad \text{for } 0 < |x| < 1.$$

Finally, we notice from the series for $g(x)$ that $g(0) = 0$.

5.3 Partial Fractions

Sometimes it does not seem easy to determine a power series for a given function $f(x)$. For example, if we try to find the Taylor Series at 0 for the function

$$f(x) = \frac{x}{(1-x)(2-x)},$$

by calculating $f^{(n)}(x)$ by direct differentiation and then writing

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(0)}{n!} x^n,$$

the algebra would very quickly get out of hand! The technique called *partial fractions* often helps us to determine such Taylor Series.

If we are given a rational function f, that is, a function of the form

$$f(x) = \frac{p(x)}{q(x)}$$

where both $p(x)$ and $q(x)$ are polynomials in x, it is useful both here and in Section 6.2 to be able to write $f(x)$ as a sum of fractions where the denominators in each case are as simple as possible. For our purposes, it is enough to consider only those original denominators $q(x)$ that can be written as a product of (linear) factors of the form $(x - \alpha)$ (for some real α) and of non-repeated (quadratic) factors of the form $x^2 + \alpha x + \beta$ (for some real α, β) that cannot be further factorized.

The first thing to do is to divide $p(x)$ by $q(x)$, by long division if necessary, to get the degree of the numerator *less* than the degree of the denominator. Suppose that the fraction we are left with is then $p_1(x)/q(x)$ for some (new) polynomial $p_1(x)$. Now factorize $q(x)$.

It can then be proved that $p_1(x)/q(x)$ can be written in a *unique* way as a sum of the following types of terms:

In all our applications of partial fractions this uniqueness property plays a vital role.

(a) corresponding to each non-repeated factor $(x - \alpha)$ of $q(x)$, a term of the form

$$\frac{C}{x - \alpha}$$

for some constant C;

(b) corresponding to each repeated factor $(x - \alpha)^r$ of $q(x)$, for given r, a collection of terms of the form

$$\frac{C_1}{x - \alpha} + \frac{C_2}{(x - \alpha)^2} + \ldots + \frac{C_r}{(x - \alpha)^r}$$

for some constants $C_1, C_2, \ldots, C_r$;

(c) corresponding to each (non-repeated) factor $x^2 + \alpha x + \beta$ of $q(x)$, that cannot be further factorized, a term of the form

$$\frac{C_1 x + C_2}{x^2 + \alpha x + \beta}$$

for some constants C_1, C_2.

This process of writing $\frac{p(x)}{q(x)}$ in terms of simpler fractions is called the **technique of partial fractions**. The above general discussion sounds rather formidable; so let us see how genuinely easy it is to apply in practice!

Example

Express $f(x) = \dfrac{x}{(1 - x)(2 - x)}$ in partial fractions.

Solution

Notice first that this is a *proper fraction*; this is, that the degree of the numerator is already less than that of the denominator. The discussion above tells us that constants C_1 and C_2 must exist such that

$$\frac{x}{(1 - x)(2 - x)} = \frac{C_1}{1 - x} + \frac{C_2}{2 - x} \qquad (1)$$

for all values of x.

Sometimes it is possible at this stage to guess the values of C_1 and C_2. An easier method, though, is to multiply up both sides of the identity (1) by the denominator of $f(x)$, to eliminate the fractions. This gives that

$$x = C_1(2 - x) + C_2(1 - x), \qquad (2)$$

which is certainly much easier to deal with.

Putting $x = 1$ in (2), we deduce that $1 = C_1$; and putting $x = 2$ in (2) gives $2 = -C_2$. Hence

$$\frac{x}{(1-x)(2-x)} = \frac{1}{1-x} - \frac{2}{2-x}.$$

Notice that you can always check your answer by bringing the terms on the right to a common denominator.

Example

Express $f(x) = \dfrac{x}{(1-x)(1+x)^2}$ in partial fractions.

Solution

Notice that again f is already a proper fraction. Our previous discussion shows that constants C_1, C_2, C_3 must exist such that

$$\frac{x}{(1-x)(1+x)^2} = \frac{C_1}{1-x} + \frac{C_2}{1+x} + \frac{C_3}{(1+x)^2}$$

so that

$$x = C_1(1+x)^2 + C_2(1-x)(1+x) + C_3(1-x). \qquad (3)$$

Putting $x = 1$ in (3) gives $1 = 4C_1$ or $C_1 = \frac{1}{4}$; and putting $x = -1$ in (3) gives $-1 = 2C_3$, so that $C_3 = -\frac{1}{2}$.

Notice that here we are using the uniqueness of C_1, C_2 and C_3—our particular choice of x is immaterial.

To find C_2, we substitute a value of x different from ± 1 into (3). Putting $x = 0$, for example, in (3) gives

$$0 = C_1 + C_2 + C_3,$$

or

$$C_2 = -C_1 - C_3 = \tfrac{1}{4}.$$

It follows that $f(x)$ can be written in the following partial fraction form:

$$f(x) = \frac{\frac{1}{4}}{1-x} + \frac{\frac{1}{4}}{1+x} - \frac{\frac{1}{2}}{(1+x)^2}.$$

Example

Express $f(x) = \dfrac{2x+1}{(1+x^2)(2-x)}$ in partial fractions,

and hence determine the Taylor Series for f at 0. Where is this series convergent?

Solution

Again, $f(x)$ is already a proper fraction. Hence, from our general discussion of partial fractions, we know that there must exist constants C_1, C_2 and C_3 such that

$$\frac{2x+1}{(1+x^2)(2-x)} = \frac{C_1 x + C_2}{1+x^2} + \frac{C_3}{2-x}.$$

Multiplying both sides by $(1+x^2)(2-x)$, we see that

$$2x + 1 = (C_1 x + C_2)(2-x) + C_3(1+x^2). \qquad (4)$$

Putting $x = 2$ in (4) gives

$$5 = 5C_3 \quad \text{so that } C_3 = 1.$$

To find C_1 and C_2 we substitute into (4) values of x different from 2. For example, putting $C_3 = 1$ in (4) and taking $x = 0, 1$ in turn, we get a pair of simultaneous equations for C_1 and C_2:

$$1 = 2C_2 + 1, \quad 3 = (C_1 + C_2) + 2.$$

It follows that $C_2 = 0$ and $C_1 = 1$. Hence the partial fractions form of $f(x)$ is

$$f(x) = \frac{x}{1+x^2} + \frac{1}{2-x}. \qquad (5)$$

Using the Geometric Series, we know that

$$\frac{1}{1+x^2} = 1 - x^2 + x^4 - \ldots + (-1)^n x^{2n} + \ldots \qquad (6)$$

for $|x^2| < 1$, and so for $|x| < 1$. Also,

$$\begin{aligned}\frac{1}{2-x} &= \frac{1}{2}\cdot\frac{1}{1-\frac{1}{2}x}\\ &= \frac{1}{2}\left\{1 + \frac{x}{2} + \left(\frac{x}{2}\right)^2 + \ldots + \left(\frac{x}{2}\right)^n + \ldots\right\} \qquad (7)\end{aligned}$$

for $\left|\frac{x}{2}\right| < 1$ and so for $|x| < 2$.

Combining (5), (6) and (7), we see that the Taylor Series for f at 0 is (expressed a little untidily!)

Notice that here we are using the uniqueness of Taylor coefficients, as discussed on page 53.

$$\begin{aligned}f(x) = x - x^3 + x^5 + \ldots + (-1)^n x^{2n+1} + \ldots\\ + \frac{1}{2}\left\{1 + \frac{x}{2} + \left(\frac{x}{2}\right)^2 + \ldots + \left(\frac{x}{2}\right)^n + \ldots\right\},\end{aligned}$$

and that this series converges for $|x| < 1$.

Self-assessment Question

Express $f(x) = \dfrac{2x}{(1+x^2)(1+x)^2}$ in partial fractions form, and hence determine the Taylor Series for f at 0. Where is this series convergent?

Solution

As the fraction is already a proper fraction, we know that there must exist constants C_1, C_2, C_3, C_4 such that

$$\frac{2x}{(1+x^2)(1+x)^2} = \frac{C_1 x + C_2}{1+x^2} + \frac{C_3}{1+x} + \frac{C_4}{(1+x)^2}. \qquad (8)$$

Multiplying both sides by $(1+x^2)(1+x)^2$ gives

$$2x = (C_1 x + C_2)(1+x)^2 + C_3(1+x)(1+x^2) + C_4(1+x^2). \qquad (9)$$

Putting $x = -1$ gives

$$-2 = 2C_4 \quad \text{so that } C_4 = -1.$$

To find the values of C_1, C_2, C_3 in (9), we substitute into (9) values of x different from -1. For example, putting $C_4 = -1$ in (9) and taking $x = 0, 1, 2$ in turn, we get

$$\begin{aligned}0 &= C_2 + C_3 - 1,\\ 2 &= 4(C_1 + C_2) + 4C_3 - 2,\\ 4 &= 9(2C_1 + C_2) + 15C_3 - 5;\end{aligned}$$

rearranging these simultaneous equations in the form

$$C_2 + C_3 = 1, \quad C_1 + C_2 + C_3 = 1 \quad \text{and } 18C_1 + 9C_2 + 15C_3 = 9,$$

it is easy to see that

$$C_1 = C_3 = 0, \; C_2 = 1.$$

Hence the partial fractions form of $f(x)$ is

$$f(x) = \frac{1}{1+x^2} - \frac{1}{(1+x)^2}.$$

Applying the Geometric Series formula to the first fraction and the result $(*)$ of part (a) of the SAQ on page 59 to the second fraction, we conclude that the Taylor Series for f at 0 is

In $(*)$ we have

$$\begin{aligned}f(x) = \{1 - x^2 + x^4 \ldots + (-1)^n x^{2n} + \ldots\}\\ - \{1 - 2x + 3x^2 + \ldots + (-1)^n (n+1)x^n + \ldots\}.\end{aligned}$$

The geometric series is convergent for $|x| < 1$, and so is the series in $(*)$. Thus the above Taylor series converges for $|x| < 1$.

6 INTEGRATION

We assume that you already know the difference between a definite integral and an indefinite integral (or primitive function), the basic techniques of integration (such as integration by parts) and standard integrals, and the Fundamental Theorem of the Calculus (which, in a sense, says that integration and differentiation are inverse processes). Accordingly, we only point out in this section some key results for *M332* that you may not have seen before or may have forgotten.

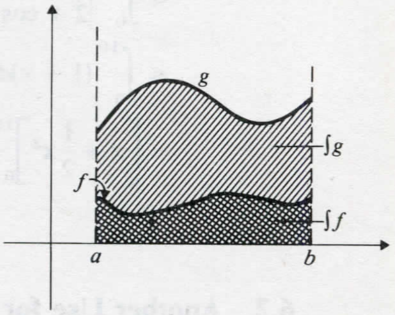
6.1 Some Inequalities for Integrals

Our first result can, somewhat loosely, be interpreted as saying that if the graph of a function f is below that of a function g, the area under the graph of f is less than that under the graph of g. More precisely

Theorem 31

Let f, g be integrable on the interval $[a,b]$ on which $f(x) \leqslant g(x)$. Then

$$\int_a^b f(x)dx \leqslant \int_a^b g(x)dx.$$



Example

Find an upper estimate for

$$I = \int_0^{1/\pi} \frac{dx}{1 + x^{10}}.$$

Solution

Since $1 + x^{10} \geqslant 1$ on $[0, \frac{1}{\pi}]$, it follows that $\frac{1}{1 + x^{10}} \leqslant 1$ on $[0, \frac{1}{\pi}]$. Hence, by Theorem 31,

$$I \leqslant \int_0^{1/\pi} 1\,dx = \frac{1}{\pi}.$$

Theorem 31 will play a key role in the proof of the Lemma on page 60 of *Unit 9* that enables us to see where functions in **C** take their maximum.

In applications, this result is often combined with our next result, which is sometimes called The Estimation Theorem or The Triangle Inequality for Integrals. (You will meet its analogue in **C** regularly in the rest of the course, starting on page 47 of *Unit 4*.)

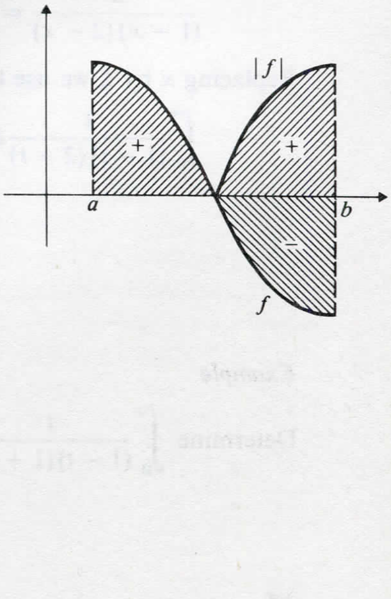
Theorem 32

Let f be integrable on the interval $[a,b]$. Then so is $|f|$, and

$$\left|\int_a^b f(x)dx\right| \leqslant \int_a^b |f(x)|dx.$$

In particular, if M is an upper bound for $|f|$ on $[a,b]$, then

$$\left|\int_a^b f(x)dx\right| \leqslant M(b-a).$$



Example

Find an upper estimate for

$$|I| = \left|\int_0^{10} \frac{1-x}{2+\cos x}dx\right|.$$

Solution

For all real x,

$$|2 + \cos x| \geqslant 1,$$

so that

$$\frac{1}{|2 + \cos x|} \leqslant 1.$$

Hence, for all $x \in [0, 10]$

$$\begin{aligned} \left|\frac{1-x}{2+\cos x}\right| &\leqslant \frac{|1-x|}{1} \\ &\leqslant 1 + x \qquad \text{(by the Triangle Inequality).} \end{aligned}$$

It follows that

$$\begin{aligned} |I| &= \left|\int_0^{10} \frac{1-x}{2+\cos x}\,dx\right| \\ &\leqslant \int_0^{10} \left|\frac{1-x}{2+\cos x}\right| dx \qquad \text{(by Theorem 32)} \\ &\leqslant \int_0^{10} (1+x)\,dx \qquad \text{(by Theorem 31)} \\ &= \left[x + \frac{1}{2}x^2\right]_0^{10} = 60 \end{aligned}$$

6.2 Another Use for Partial Fractions

In Section 5.3 we explained what partial fractions are, and how they can be sometimes used to find power series expansions of rational functions. In exactly the same way, it is often possible to use partial fractions to find the integrals of rational functions.

Example

Determine $\displaystyle\int_0^x \frac{t}{(1-t)(2-t)}\,dt$ for $x < 1$.

Solution

We saw on page 61 that

$$\frac{x}{(1-x)(2-x)} = \frac{1}{1-x} - \frac{2}{2-x}.$$

Replacing x by t, we use this to deduce that

$$\begin{aligned} \int_0^x \frac{t}{(1-t)(2-t)}\,dt &= \int_0^x \left(\frac{1}{1-t} - \frac{2}{2-t}\right) dt \\ &= \Big[-\log_e(1-t) + 2\log_e(2-t)\Big]_0^x \\ &= \log_e\left\{\frac{(2-x)^2}{4(1-x)}\right\}. \end{aligned}$$

Example

Determine $\displaystyle\int_0^x \frac{t}{(1-t)(1+t)^2}\,dt$ for $-1 < x < 1$.

Solution

We saw on page 61 that

$$\frac{x}{(1-x)(1+x)^2} = \frac{\frac{1}{4}}{1-x} + \frac{\frac{1}{4}}{1+x} - \frac{\frac{1}{2}}{(1+x)^2}.$$

It follows that

$$\begin{aligned}\int_0^x \frac{t}{(1-t)(1+t)^2}\,dt &= \int_0^x \left\{\frac{\frac{1}{4}}{1-t} + \frac{\frac{1}{4}}{1+t} - \frac{\frac{1}{2}}{(1+t)^2}\right\} dt\\ &= \left[-\tfrac{1}{4}\log_e(1-t) + \tfrac{1}{4}\log_e(1+t) + \frac{\frac{1}{2}}{1+t}\right]_0^x\\ &= \tfrac{1}{4}\log_e\left(\frac{1+x}{1-x}\right) + \frac{\frac{1}{2}}{1+x} - \frac{1}{2}.\end{aligned}$$

Self-assessment Question

Determine $\displaystyle\int_0^x \frac{2t+1}{(1+t^2)(2-t)}\,dt$ for $x < 2$.

Solution

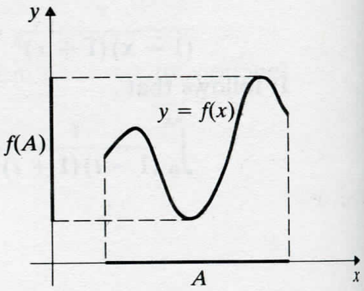
Using the partial fractions expansion

$$\frac{2x+1}{(1+x^2)(2-x)} = \frac{x}{1+x^2} + \frac{1}{2-x}$$

on page 61, we deduce that

$$\begin{aligned}\int_0^x \frac{2t+1}{(1+t^2)(2-t)}\,dt &= \int_0^x \left(\frac{t}{1+t^2} + \frac{1}{2-t}\right) dt\\ &= \left[\tfrac{1}{2}\log_e(1+t^2) - \log_e(2-t)\right]_0^x\\ &= \tfrac{1}{2}\log_e(1+x^2) - \log_e(2-x) + \log_e 2\\ &= \log_e\left(\frac{2(1+x^2)^{1/2}}{2-x}\right).\end{aligned}$$

7 PARTIAL DERIVATIVES



The *graph* of a function f of one real variable is the set of points $(x, f(x))$ in the plane $\mathbf{R}^2$ for which x lies in the domain of f—that is, some subset A of $\mathbf{R}$. The graph of f is useful when dealing with a continuous or differentiable function.

However, in many branches of mathematics, the functions in which we are interested depend on more than one variable. In *M332* we are interested in functions of *two* real variables—that is, functions f whose domain is some subset A of $\mathbf{R}^2$:

$$\begin{aligned} f\colon A &\longrightarrow \mathbf{R} \\ (x, y) &\longmapsto f(x, y). \end{aligned}$$

Example

If $f(x, y) = 3x^2 - 2y^3$, what is the value of $f(2, 3)$?

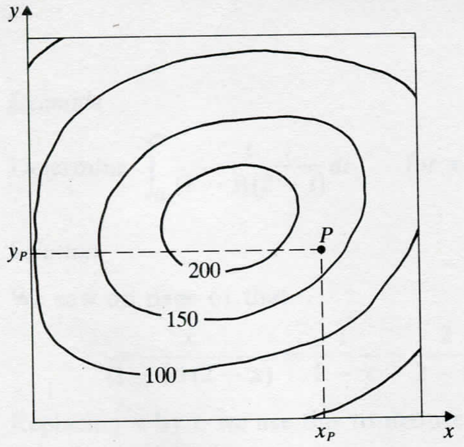
Solution

To get this value, we substitute 2 for x and 3 for y in the expression for $f(x, y)$; this gives

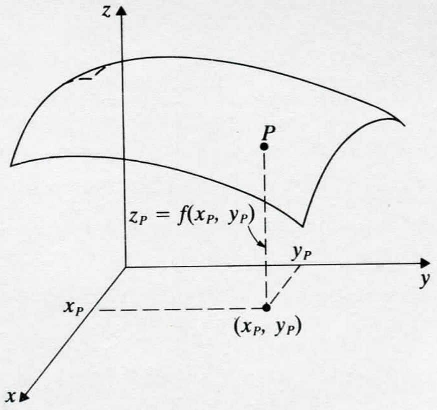
$$f(2, 3) = 3(2)^2 - 2(3)^3 = 12 - 54 = -42.$$

For functions f of two variables we may define the continuity of f in a way similar to that for functions of one variable. However, there is no simple analogue for differentiation. In this section we explain what replaces the derivative of a function of one variable, the so-called *partial derivatives.*

To motivate our definition of partial derivatives, we consider a map of a region of countryside that contains a hill (see Figure (a)). In order to refer to the position of any point P in this region, we draw fixed x– and y– coordinate axes on the map at right angles to each other, and specify the location (x_P, y_P) of the point P on the map by giving its distance from each axis.



(a) map of a hill



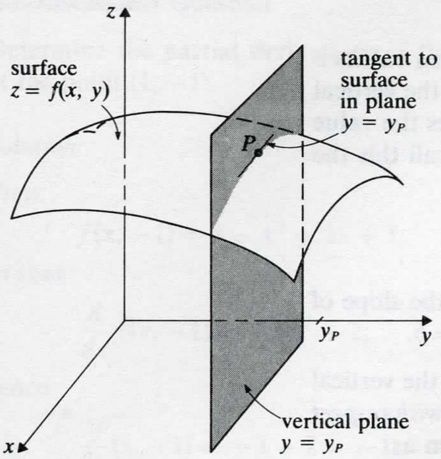
(b) surface $z = f(x,y)$

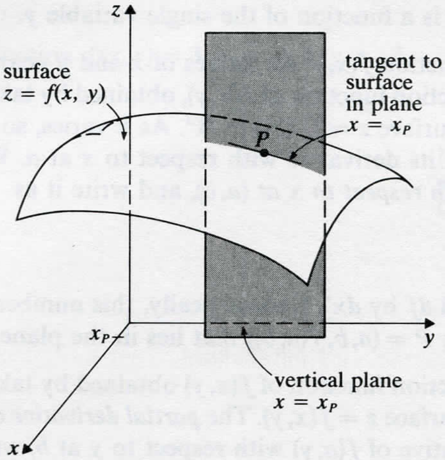
Maps of the countryside usually show contour lines—that is, lines of constant height. Using these, we can specify the location of P more precisely by assigning to the point (x_P, y_P) on the map a third number, z_P, which is the height of P above sea-level. Corresponding to each choice of coordinates (x, y), there is a corresponding number $z = f(x, y)$, and we say that the function f defined in this way is a function of (that is, is completely determined by) the two variables x and y. Given a function $f(x, y)$, we say that the set of points $(x, y, f(x, y))$ in $\mathbf{R}^3$ is the *surface with equation* $z = f(x, y)$.

A man (of negligible size!) standing on the hill at the point P who looks south from P (that is, in the direction of the vertical plane $y = y_P$ through P) where

P is the particular point shown below, will see that the hill has a zero slope in this direction; in other words, the tangent at P to the surface that lies in the direction of the plane $y = y_P$ has zero slope.

On the other hand, if the man looks east from P (that is, in the direction of the vertical plane $x = x_P$ through P) he will see that the hill has a negative slope in this direction; in other words, the tangent at P to the surface that lies in the direction of the plane $x = x_P$ has negative slope.





We shall define the slopes of tangent lines to a surface $z = f(x, y)$ that lie in particular vertical planes to be the *partial derivatives* of f at P. This is analogous to our definition of the *derivative* of a function f of one variable as the slope of the tangent to the corresponding graph $y = f(x)$.

In general, a *section* of a surface S with equation $z = f(x, y)$ is the intersection of S with a vertical plane. The two most convenient vertical planes for our purpose through a point $(a, b, f(a, b))$ on S are the planes through the point that are parallel to the (x, z)–plane and the (y, z)–plane; these have equations $y = b$ and $x = a$, respectively.

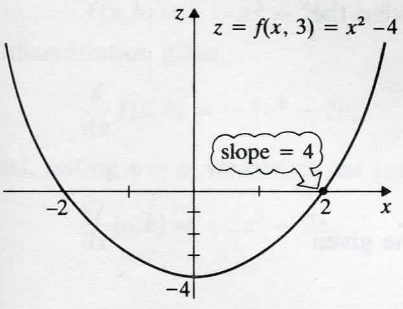
Here we use a, b rather than x_P, y_P.

Example

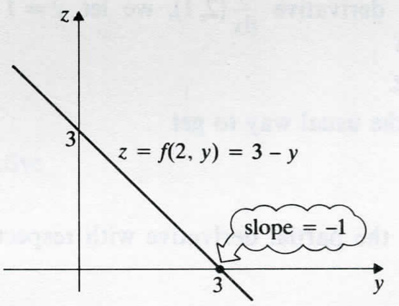
Sketch the sections of the surface S with equation $z = x^2 - y - 1$ obtained as the intersections of S with the planes $y = 3$ and $x = 2$.

Solution

Put $f(x, y) = x^2 - y - 1$. Substituting $y = 3$ and $x = 2$ in the formula for $f(x, y)$ and sketching the graphs of the resulting functions $x \longmapsto x^2 - 4$ and $y \longmapsto 3 - y$, we obtain the following sections.



intersection of S with plane $y = 3$



intersection of S with plane $x = 2$

Thus a man standing on the surface S at the point $(2, 3, 0)$ would see its section by the plane $y = 3$ as a parabola (whose slope is 4 at the point where he is standing), and its section by the plane $x = 2$ as a line with slope -1.

We say that the graphs

$$z = f(x, 3) \qquad \text{and} \qquad z = f(2, y)$$

are the graphs of the *section functions*

$$x \longmapsto f(x, 3) \qquad \text{and} \qquad y \longmapsto f(2, y).$$

Thus $f(x, 3)$ is obtained from the function $f(x, y) = x^2 - y - 1$ of two variables by holding y at the fixed value 3, so that $f(x, 3) = x^2 - 4$ is a function of the single variable x. Similarly, $f(2, y)$ is obtained by holding x at the fixed value 2, so that $f(2, y) = 3 - y$ is a function of the single variable y.

In general, consider a function $f(x, y)$ for values of x and y near given values a and b. Then $f(x, b)$ is a section function of $f(x, y)$, obtained by taking the vertical slice $y = b$ through the surface $z = f(x, y)$ in $\mathbf{R}^3$. As x varies, so does the value of $f(x, b)$; so we can find its derivative with respect to x at a. We call this the *partial derivative of f with respect to x at (a, b)*, and write it as

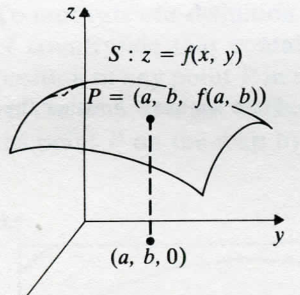
$$\frac{\partial f}{\partial x}(a, b),$$

which we read as 'partial df by dx'. Geometrically, this number is the slope of the tangent to S through $P = (a, b, f(a, b))$ that lies in the plane $y = b$.

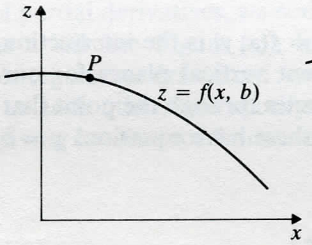
Similarly, $f(a, y)$ is the section function of $f(x, y)$ obtained by taking the vertical slice $x = a$ through the surface $z = f(x, y)$. The *partial derivative of f with respect to y at (a, b)* is the derivative of $f(a, y)$ with respect to y at b, written as

$$\frac{\partial f}{\partial y}(a, b).$$

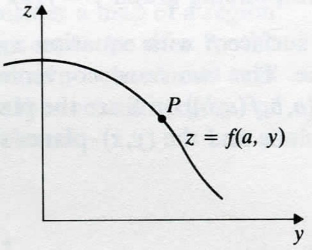
It is the slope of the tangent to S through P that lies in the plane $x = a$.



surface in $\mathbf{R}^3$



plane $y = b$



plane $x = a$

Example

Determine the partial derivatives of the function $f(x, y) = 3x^2 - 2y^3$ at the point $(2, 1)$.

Solution

To calculate the partial derivative $\frac{\partial f}{\partial x}(2, 1)$, we let $y = 1$ and consider the resulting section function

$$f(x, 1) = 3x^2 - 2.$$

We now differentiate in the usual way to get

$$\frac{d}{dx} f(x, 1) = 6x.$$

Setting $x = 2$, we obtain the partial derivative with respect to x of the given function at $(2, 1)$:

$$\frac{\partial f}{\partial x}(2, 1) = 6 \cdot 2 = 12.$$

Similarly, to calculate $\frac{\partial f}{\partial y}(2, 1)$, we let $x = 2$ and consider the resulting section function

$$f(2, y) = 12 - 2y^3.$$

We now differentiate in the usual way to get

$$\frac{d}{dy}f(2, y) = -6y^2.$$

Setting $y = 1$, we obtain the partial derivative

$$\frac{\partial f}{\partial y}(2, 1) = -6.$$

Self-assessment Question

Determine the partial derivatives of the function $f(x, y) = 3 - x^3 - 2xy + y^2$ at the point $(1, -1)$.

Solution

In this Solution we omit all unnecessary explanations; thus a solution of the type given here is quite adequate.

First,

$$f(x, -1) = 3 - x^3 + 2x + 1,$$

so that

$$\frac{d}{dx}f(x, -1) = -3x^2 + 2;$$

hence

$$\frac{\partial f}{\partial x}(1, -1) = -3 + 2 = -1.$$

Next,

$$f(1, y) = 2 - 2y + y^2,$$

so that

$$\frac{d}{dy}f(1, y) = -2 + 2y;$$

hence

$$\frac{\partial f}{\partial y}(1, -1) = -2 - 2 = -4.$$

We can calculate the partial derivatives of $f(x, y)$ at an arbitrary point (a, b) in just the same way as for the point $(1, -1)$.

Example

For the function $f(x, y) = 3 - x^3 - 2xy + y^2$, determine $\frac{\partial f}{\partial x}(a, b)$.

Solution

We first fix y at the value b and consider the section function

$$f(x, b) = 3 - x^3 - 2bx + b^2.$$

Differentiation gives

$$\frac{d}{dx}f(x, b) = -3x^2 - 2b;$$

and, setting $x = a$, we obtain the partial derivative

$$\frac{\partial f}{\partial x}(a, b) = -3a^2 - 2b.$$

Self-assessment Question

For the same function, determine $\frac{\partial f}{\partial y}(a, b)$.

Solution

We first fix x at the value a and consider the section function

$$f(a, y) = 3 - a^3 - 2ay + y^2.$$

Differentiation gives

$$\frac{d}{dy} f(a, y) = -2a + 2y;$$

and, setting $y = b$, we obtain the partial derivative

$$\frac{\partial f}{\partial y}(a, b) = -2a + 2b.$$

Because a and b are quite general, we could just as well have taken x and y in place of a and b in the above Example. We would then have defined the function $\frac{\partial f}{\partial x}(x, y)$, abbreviated to $\frac{\partial f}{\partial x}$, as a function of x and y. To calculate $\frac{\partial f}{\partial x}$ we just regard y as a constant, and differentiate f (partially!) with respect to x. We can define and calculate $\frac{\partial f}{\partial y}$ in a similar way, by fixing x and differentiating with respect to y.

Example

Determine the partial derivatives $\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}$ of the function $f(x, y) = x - y^2 + x^2y^3$.

Solution

First we differentiate f with respect to x, while treating y as a constant. This gives

$$\frac{\partial f}{\partial x}(x, y) = 1 + 2xy^3.$$

Next we differentiate f with respect to y, while treating x as a constant. This gives

$$\frac{\partial f}{\partial y}(x, y) = -2y + 3x^2y^2.$$

Self-assessment Question

Determine the partial derivatives $\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}$ of the function

$$f(x, y) = 2 + 3x - y + 4xy - 2x^3y^2.$$

Solution

First we differentiate f with respect to x, while treating y as a constant. This gives

$$\frac{\partial f}{\partial x}(x, y) = 3 + 4y - 6x^2y^2.$$

Next we differentiate f with respect to y, while treating x as a constant. This gives

$$\frac{\partial f}{\partial y}(x, y) = -1 + 4x - 4x^3y.$$

Self-assessment Question

For the function $f(x, y) = x^3 - 3x^2y^4 + y^5$, determine $\frac{\partial f}{\partial x}$, $\frac{\partial f}{\partial x}(1, 1)$, $\frac{\partial f}{\partial y}$ and $\frac{\partial f}{\partial y}(1, 1)$.

Solution

First we differentiate f with respect to x, while treating y as a constant. This gives

$$\frac{\partial f}{\partial x}(x, y) = 3x^2 - 6xy^4.$$

Substituting $x = 1$, $y = 1$, we obtain

$$\frac{\partial f}{\partial x}(1, 1) = 3 - 6 = -3.$$

Next we differentiate f with respect to y, while treating x as a constant. This gives

$$\frac{\partial f}{\partial y}(x, y) = -12x^2y^3 + 5y^4.$$

Substituting $x = 1$, $y = 1$, we obtain

$$\frac{\partial f}{\partial y}(1, 1) = -12 + 5 = -7.$$

Produced in Great Britain by Speedlith Photo Litho Limited, Manchester M32 0JT

0 335 05565 6